

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 15:57:32 ; Search time 3207.21 Seconds  
(without alignments)  
9775.823 Million cell updates/sec

Title: US-09-909-207-1  
Perfect score: 663  
Sequence: 1 CAAATCGTCACCGACAATTC.....TAACCTTGGATAAAACAAT 663

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba.\*  
2: gb\_hgt.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

A48222 Sequence 1 from Patent EP0898667. 663 bp DNA linear PAT 07-MAR-1997

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	663	100.0	663	6	A48222 Sequence 1
2	663	100.0	663	6	A48223 Sequence 2
3	663	100.0	663	6	AR193049 Sequence
4	663	100.0	663	6	AR193050 Sequence
5	663	100.0	744	6	A48225 Sequence 4
6	663	100.0	744	6	A48226 Sequence 5
7	663	100.0	744	6	AR193051 Sequence
8	663	100.0	744	6	AR193052 Sequence
9	663	100.0	1513	6	A48231 Sequence 10
10	663	100.0	1513	6	A48232 Sequence 11
11	663	100.0	1513	6	AR193055 Sequence
12	663	100.0	1513	6	AR193056 Sequence
13	634.2	95.7	744	6	A68006 Sequence 1
14	634.2	95.7	744	6	AR163110 Sequence
15	624.6	94.2	744	6	A45313 Sequence 18
16	624.6	94.2	744	6	AR117325 Sequence
17	507.8	76.6	744	6	A68016 Sequence 11
18	507.8	76.6	744	6	AR163117 Sequence
19	355.2	53.6	1531	1	BSXNY

20	355.2	53.6	2246	1	AB029319	Bacillus
21	280.4	42.3	1454	1	CLOXYNB	M31726 Clostridium
22	278.2	42.0	4128	1	AF047761	AB047761 Clostridi
23	276.6	41.7	5774	1	AB010958	AB010958 Clostridi
24	274.8	41.4	3493	1	CSF508403	AJ508403 Clostridi
25	268.2	40.5	1070	1	BPXYNA	X06660 Bacillus pu
26	266.6	40.2	687	1	AY526092	AY526092 Bacillus
27	263.4	39.7	1011	1	AF490981	AF490981 Bacillus
28	258.6	39.0	1789	1	AF326785	AF326785 Bacillus
29	255.4	38.5	2364	1	CLOXYNA	D13325 Clostridium
30	255.4	38.5	2364	6	AR203507	AR203507 Sequence
31	255.4	38.5	2364	6	AR373315	AR373315 Sequence
32	255.2	38.5	1022	6	A42251	A42251 Sequence 1
33	255.2	38.5	1022	6	A42285	A42285 Sequence 35
34	255.2	38.5	1022	6	AR127019	AR127019 Sequence
35	255.2	38.5	1022	6	AR127049	AR127049 Sequence
36	255.2	38.5	1022	6	AR220023	AR220023 Sequence
37	255.2	38.5	1022	6	AR220053	AR220053 Sequence
38	255.2	38.5	1022	6	AR221306	AR221306 Sequence
39	255.2	38.5	1022	6	AR221336	AR221336 Sequence
40	255	38.5	600	6	A42276	A42276 Sequence 26
41	255	38.5	600	6	A42284	A42284 Sequence 34
42	255	38.5	600	6	AR127041	AR127041 Sequence
43	255	38.5	600	6	AR127048	AR127048 Sequence
44	255	38.5	600	6	AR220045	AR220045 Sequence
45	255	38.5	600	6	AR220052	AR220052 Sequence

## ALIGNMENTS

RESULT 1

A48222

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .663

/organism="unidentified"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32644"

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

0;

100.0%; Score 663; DB 6; Length 663;

100.0%; Pred. No. 2.1e-169;

0; Mismatches 0; Indels 0; Gaps 0;

1 CAAATCGTCACCGACAATTC

1 CAAATCGTCACCGACAATTC

61 CAAATCGTCACCGACAATTC

61 CAAATCGTCACCGACAATTC

QY	121	AACAAATGTTAAACAATATTTCCGTAAGGTAAAAAATTCATATGAAACACAAACACAC	180
Db	121	AACAAATGTTAAACAATATTTCCGTAAGGTAAAAAATTCATATGAAACACAAACACAC	180
QY	181	CAACAAGTTGGTAAACATGTCATAACTACGAGGCAACCTTCCAAACCAATGTTAATGCG	240
Db	181	CAACAAGTTGGTAAACATGTCATAACTACGAGGCAACCTTCCAAACCAATGTTAATGCG	240
QY	241	TATTTATGCGTCTATGTTGGTGGACTGTTGACCTCTTGTGGAATATTTATTTGTCGACAGT	300
Db	241	TATTTATGCGTCTATGTTGGTGGACTGTTGACCTCTTGTGGAATATTTATTTGTCGACAGT	300
QY	301	TGGGGCAACTGGCGTCCACAGGCAACCGCTTAAGGGGACCATCTGTTGATGAGGA	360
Db	301	TGGGGCAACTGGCGTCCACAGGCAACCGCTTAAGGGGACCATCTGTTGATGAGGA	360
QY	361	ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCACA	420
Db	361	ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCACA	420
QY	421	TTTAAACAATATTTGAGTGTTCGAAGATCGAAACGACGAGTGGCAGCATTTCTGTCAGC	480
Db	421	TTTAAACAATATTTGAGTGTTCGAAGATCGAAACGACGAGTGGCAGCATTTCTGTCAGC	480
QY	481	AACCACTTTAGAGCGTGGGAAACTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCACA	540
Db	481	AACCACTTTAGAGCGTGGGAAACTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCACA	540
QY	541	CTTACTGTAGAGCGTCTCAACTATTTAGTAATGACGAGCATAACTTTGGATAAAAAC	600
Db	541	CTTACTGTAGAGCGTCTCAACTATTTAGTAATGACGAGCATAACTTTGGATAAAAAC	600
QY	601	ATTAAACGTTAAACCTCTCTCAACTATTAGTAATGACGAGCATAACTTTGGATAAAAAC	660
Db	601	ATTAAACGTTAAACCTCTCTCAACTATTAGTAATGACGAGCATAACTTTGGATAAAAAC	660
QY	661	AAT 663	
Db	661	AAT 663	
RESULT 2			
A48223	Sequence 2 from Patent EP0698667.		
LOCUS	A48223.1 GI:2302070		
DEFINITION	unidentified		
ACCESSION	unidentified		
VERSION	unidentified		
KEYWORDS	unclassified.		
SOURCE	De,B.E., Lahave,A., Ledoux,P. and Detroz,R.		
ORGANISM	Xylanase, microorganisms for its production, DNA molecules, process of preparation and use thereof		
REFERENCE	Patent: EP 0698667-A 2 28-FEB-1996;		
AUTHORS	SOLVAY (BE)		
TITLE	Other publication BE 1008751 960702		
JOURNAL	Other publication BE 1008570 960604		
COMMENT	Other publication BR 9503454 960305		
	Other publication JP 8092284 960409		
	Other publication FI 953578 960127		
	Other publication CA 2154628 960127		
	Other publication AU 2508695 960208.		
FEATURES			
source	1. .663		
	/organism="unidentified"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32644"		
CDS	1..663		
	/note="unnamed protein product; Protein sequence is in conflict with the conceptual translation"		
	/codon_start=1		

/protein_id="CAA03091.1"			
/db_xref="GI:2302070"			
/translation="Q1VTDNSIGNHDGYDFEWMKDSGSGTMTLHNHGGTFSQAWNVDNNGNINLFRKGKFNETHQOQVGNMSINYGNFOPNGNAYLCVYQHTWPLVEYVYIDVSNWNRPFPGATPGKGTITVDGGTYDIYETLRVNQPSIKNGIAIFKOYWSVRKSKTSIGTISVSNHFRAWENLGNMNGMYEVALTVBEGYQSGSANVYNTLRNGNPLUSTISNDESITLIDKNN"			
mat_peptide	1..663		
/product="unnamed"			
ORIGIN			
Query Match 100.0%; Score 663; DB 6; Length 663;			
Best Local Similarity 100.0%; Pred. No. 2.1e-169;			
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CAAAATCGTCAACGACAAATTCATTTGGCAACACGATGGCTATGATTAATGATTTTGGAAA	60
Db	1	CAAAATCGTCAACGACAAATTCATTTGGCAACACGATGGCTATGATTAATGATTTTGGAAA	60
QY	61	GATAGCGGTGGCTCTGGGCAATGATTTCTCAATCATGCGGTACGTTCAAGTGCCTCAATGG	120
Db	61	GATAGCGGTGGCTCTGGGCAATGATTTCTCAATCATGCGGTACGTTCAAGTGCCTCAATGG	120
QY	121	AACAATGTTAAACAATATTTCCGTAAAGGTAAAAAATTCATGAAACACAAACACAC	180
Db	121	AACAATGTTAAACAATATTTCCGTAAAGGTAAAAAATTCATGAAACACAAACACAC	180
QY	181	CAACAAGTTGGTAAACATGTCATAACTACGAGGCAACCTTCCAAACCAATGTTAATGCG	240
Db	181	CAACAAGTTGGTAAACATGTCATAACTACGAGGCAACCTTCCAAACCAATGTTAATGCG	240
QY	241	TATTTATGCGTCTATGTTGGTGGACTGTTGACCTCTTGTGGAATATTTATTTGTCGACAGT	300
Db	241	TATTTATGCGTCTATGTTGGTGGACTGTTGACCTCTTGTGGAATATTTATTTGTCGACAGT	300
QY	301	TGGGGCAACTGGCGTCCACAGGCAACCGCTTAAGGGGACCATCTGTTGATGAGGA	360
Db	301	TGGGGCAACTGGCGTCCACAGGCAACCGCTTAAGGGGACCATCTGTTGATGAGGA	360
QY	361	ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCACA	420
Db	361	ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCACA	420
QY	421	TTTAAACAATATTTGAGTGTTCGAAGATCGAAACGACGAGTGGCAGCATTTCTGTCAGC	480
Db	421	TTTAAACAATATTTGAGTGTTCGAAGATCGAAACGACGAGTGGCAGCATTTCTGTCAGC	480
QY	481	AACCACTTTAGAGCGTGGGAAACTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCAG	540
Db	481	AACCACTTTAGAGCGTGGGAAACTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCAG	540
QY	541	CTTACTGTAGAGCGTCTCAACTATTTAGTAATGACGAGCATAACTTTGGATAAAAAC	600
Db	541	CTTACTGTAGAGCGTCTCAACTATTTAGTAATGACGAGCATAACTTTGGATAAAAAC	600
QY	601	ATTAAACGTTAAACCTCTCTCAACTATTAGTAATGACGAGCATAACTTTGGATAAAAAC	660
Db	601	ATTAAACGTTAAACCTCTCTCAACTATTAGTAATGACGAGCATAACTTTGGATAAAAAC	660
QY	661	AAT 663	
Db	661	AAT 663	
RESULT 3			
AR193049	Sequence 1 from patent US 6346407.		
LOCUS	AR193049.1 GI:20239014		
DEFINITION	Unknown.		
ACCESSION	Unknown.		
VERSION	Unknown.		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		

Unclassified.  
REFERENCE 1 (bases 1 to 663)  
AUTHORS De Buyl, E., Lahaye, A., Ledoux, P. and Detroz, R.  
TITLE Xylanase, microorganisms producing it, DNA molecules, methods for preparing this xylanase and uses of the latter  
JOURNAL Patent: US 6346407-A 1 12-FEB-2002;  
FEATURES Location/Qualifiers  
source 1..663  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 663; DB 6; Length 663;  
Best Local Similarity 100.0%; Pred. No. 2.1e-169;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAATCGTCACGACAAATCCATTTGGCAACACGATGGCTATGATATCAATTTTGGAAA 60  
DB 1 CAAATCGTCACGACAAATCCATTTGGCAACACGATGGCTATGATATCAATTTTGGAAA 60

QY 61 GATAGCGGTGGCTCTGGGACAAATGATCTCAATCATGGCGGTACGTTCAAGTCCCAATGG 120  
DB 61 GATAGCGGTGGCTCTGGGACAAATGATCTCAATCATGGCGGTACGTTCAAGTCCCAATGG 120

QY 121 AACATGTTAAACAACATATTTATTCGTTAAAGGTAAAGGTTCAATGAAAACACAAACACAC 180  
DB 121 AACATGTTAAACAACATATTTATTCGTTAAAGGTAAAGGTTCAATGAAAACACAAACACAC 180

QY 181 CAAATGTTGTTAAACATGTCATATAAATACGAGGCAACCTTCCAAACCAATGGTATGGC 240  
DB 181 CAAATGTTGTTAAACATGTCATATAAATACGAGGCAACCTTCCAAACCAATGGTATGGC 240

QY 241 TATTTATGCGTCTATGGTGGGACTGTTGACCTCTTGTGCAATATTTATTTGCGACAGT 300  
DB 241 TATTTATGCGTCTATGGTGGGACTGTTGACCTCTTGTGCAATATTTATTTGCGACAGT 300

QY 301 TGGGGCAACTGGCGTCCACGAGCAACGCCCTAAGGGGACCATCACTGTTGATGGAGGA 360  
DB 301 TGGGGCAACTGGCGTCCACGAGCAACGCCCTAAGGGGACCATCACTGTTGATGGAGGA 360

QY 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATGGCCACA 420  
DB 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATGGCCACA 420

QY 421 TTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGATTTCTGTGAGC 480  
DB 421 TTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGATTTCTGTGAGC 480

QY 481 AACCACTTTAGAGCGTGGGAAAACCTTAGGATGAATATGGGAAAATGTATGAAGTCGCG 540  
DB 481 AACCACTTTAGAGCGTGGGAAAACCTTAGGATGAATATGGGAAAATGTATGAAGTCGCG 540

QY 541 CTTACTGTAGAAGCTATCAAGATAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA 600  
DB 541 CTTACTGTAGAAGCTATCAAGATAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA 600

QY 601 ATTAACGGTAACCCCTCTCAACTATTTAGTAATGACGAGAGCATACCTTTGGATAAAAC 660  
DB 601 ATTAACGGTAACCCCTCTCAACTATTTAGTAATGACGAGAGCATACCTTTGGATAAAAC 660

QY 661 AAT 663  
DB 661 AAT 663

RESULT 4  
AR193050  
LOCUS AR193050 663 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 2 from patent US 6346407.  
ACCESSION AR193050  
VERSION AR193050.1 GI:20239015  
KEYWORDS Unknown.

ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 663)  
TITLE De Buyl, E., Lahaye, A., Ledoux, P. and Detroz, R.  
JOURNAL Xylanase, microorganisms producing it, DNA molecules, methods for preparing this xylanase and uses of the latter  
FEATURES Patent: US 6346407-A 2 12-FEB-2002;  
source Location/Qualifiers  
1..663  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 663; DB 6; Length 663;  
Best Local Similarity 100.0%; Pred. No. 2.1e-169;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAATCGTCACGACAAATCCATTTGGCAACACGATGGCTATGATATCAATTTTGGAAA 60  
DB 1 CAAATCGTCACGACAAATCCATTTGGCAACACGATGGCTATGATATCAATTTTGGAAA 60

QY 61 GATAGCGGTGGCTCTGGGACAAATGATCTCAATCATGGCGGTACGTTCAAGTCCCAATGG 120  
DB 61 GATAGCGGTGGCTCTGGGACAAATGATCTCAATCATGGCGGTACGTTCAAGTCCCAATGG 120

QY 121 AACATGTTAAACAACATATTTATTCGTTAAAGGTAAAGGTTCAATGAAAACACAAACACAC 180  
DB 121 AACATGTTAAACAACATATTTATTCGTTAAAGGTAAAGGTTCAATGAAAACACAAACACAC 180

QY 181 CAAATGTTGTTAAACATGTCATATAAATACGAGGCAACCTTCCAAACCAATGGTATGGC 240  
DB 181 CAAATGTTGTTAAACATGTCATATAAATACGAGGCAACCTTCCAAACCAATGGTATGGC 240

QY 241 TATTTATGCGTCTATGGTGGGACTGTTGACCTCTTGTGCAATATTTATTTGCGACAGT 300  
DB 241 TATTTATGCGTCTATGGTGGGACTGTTGACCTCTTGTGCAATATTTATTTGCGACAGT 300

QY 301 TGGGGCAACTGGCGTCCACGAGCAACGCCCTAAGGGGACCATCACTGTTGATGGAGGA 360  
DB 301 TGGGGCAACTGGCGTCCACGAGCAACGCCCTAAGGGGACCATCACTGTTGATGGAGGA 360

QY 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATGGCCACA 420  
DB 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATGGCCACA 420

QY 421 TTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGATTTCTGTGAGC 480  
DB 421 TTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGATTTCTGTGAGC 480

QY 481 AACCACTTTAGAGCGTGGGAAAACCTTAGGATGAATATGGGAAAATGTATGAAGTCGCG 540  
DB 481 AACCACTTTAGAGCGTGGGAAAACCTTAGGATGAATATGGGAAAATGTATGAAGTCGCG 540

QY 541 CTTACTGTAGAAGCTATCAAGATAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA 600  
DB 541 CTTACTGTAGAAGCTATCAAGATAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA 600

QY 601 ATTAACGGTAACCCCTCTCAACTATTTAGTAATGACGAGAGCATACCTTTGGATAAAAC 660  
DB 601 ATTAACGGTAACCCCTCTCAACTATTTAGTAATGACGAGAGCATACCTTTGGATAAAAC 660

QY 661 AAT 663  
DB 661 AAT 663

RESULT 5  
A48225  
LOCUS A48225 744 bp DNA linear PAT 07-MAR-1997  
DEFINITION Sequence 4 from Patent EP0698667.  
ACCESSION A48225  
VERSION A48225.1 GI:2302072  
KEYWORDS

SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 744)  
AUTHORS De,B.E., Lahaye,A., Ledoux,P. and Detroz,R.  
TITLE Xylanase, microorganisms for its production, DNA molecules, process of preparation and use thereof  
JOURNAL Patent: EP 0698667-A 4 28-FEB-1996;  
SOLVAY (BE)  
COMMENT Other publication BE 1008751 960702  
Other publication BE 1008570 960604  
Other publication BR 9503454 960305  
Other publication JP 8092284 960409  
Other publication FI 953578 960127  
Other publication CA 2154628 960127  
Other publication AU 2508695 960208.  
FEATURES  
source  
1. 744  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Query Match 100.0%; Score 663; DB 6; Length 744;  
Best Local Similarity 100.0%; Pred. No. 2.1e-169; Mismatches 0; Indels 0; Gaps 0;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAATCGTCACCGACAAATTCATTCGCAACCAACGATGGCTATGATTATGAATTTGGAAA 60  
DB 82 CAATCGTCACCGACAAATTCATTCGCAACCAACGATGGCTATGATTATGAATTTGGAAA 141  
QY 61 GATAGCGGTGGCTCTGGGACAAATGATTCATCATGATGGGGTACGTTGAGTCCCAATGG 120  
DB 142 GATAGCGGTGGCTCTGGGACAAATGATTCATCATGATGGGGTACGTTGAGTCCCAATGG 201  
QY 121 ACAATGTTAAACAATATTTTCGTAAGGTAAATAATTCATGAAACACAAACACAC 180  
DB 202 ACAATGTTAAACAATATTTTCGTAAGGTAAATAATTCATGAAACACAAACACAC 261  
QY 181 CAACAAGTTGGTAACATGTCCTAATAAATAACGAGGCAACTTCACACCAAAATGGTAATGG 240  
DB 262 CAACAAGTTGGTAACATGTCCTAATAAATAACGAGGCAACTTCACACCAAAATGGTAATGG 321  
QY 241 TATTTATGCGTCTATGGTTGGACTGTTGACCTCTGTGCGATATATTTATGTCACAGT 300  
DB 322 TATTTATGCGTCTATGGTTGGACTGTTGACCTCTGTGCGATATATTTATGTCACAGT 381  
QY 301 TGGGGCAACTGGCGTCCACGAGGACCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 360  
DB 382 TGGGGCAACTGGCGTCCACGAGGACCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 441  
QY 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCATTAAGGGGATGGCCACA 420  
DB 442 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCATTAAGGGGATGGCCACA 501  
QY 421 TTTAAACAATTTGGAGTCTCGAAGATGAAACGACGAGTGGGACGATTTCTGTGACG 480  
DB 502 TTTAAACAATTTGGAGTCTCGAAGATGAAACGACGAGTGGGACGATTTCTGTGACG 561  
QY 481 AACCACTTTAGCGGTGGGAAAATCTTAGGATGAATATGGGAAAATATGATGAAGTCGCG 540  
DB 562 AACCACTTTAGCGGTGGGAAAATCTTAGGATGAATATGGGAAAATATGATGAAGTCGCG 621  
QY 541 CTCTACTGTAGAGGCTATCAAGTAGCGGAGTGTAAATGTTATATAGCAATACACTAAGA 600  
DB 622 CTCTACTGTAGAGGCTATCAAGTAGCGGAGTGTAAATGTTATATAGCAATACACTAAGA 681  
QY 601 ATTTAAGGTAACCTCTCTCAACTATTAGTAATGACGAGGACATACTTTGGATAAAGAC 660  
DB 682 ATTTAAGGTAACCTCTCTCAACTATTAGTAATGACGAGGACATACTTTGGATAAAGAC 741  
QY 661 AAT 663  
|||

Db 742 AAT 744  
RESULT 6  
LOCUS A48226  
DEFINITION Sequence 5 from Patent EP0698667.  
ACCESSION A48226  
VERSION A48226.1 GI:2302073  
KEYWORDS unidentified  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 744)  
AUTHORS De,B.E., Lahaye,A., Ledoux,P. and Detroz,R.  
TITLE Xylanase, microorganisms for its production, DNA molecules, process of preparation and use thereof  
JOURNAL Patent: EP 0698667-A 5 28-FEB-1996;  
SOLVAY (BE)  
COMMENT Other publication BE 1008751 960702  
Other publication BE 1008570 960604  
Other publication BR 9503454 960305  
Other publication JP 8092284 960409  
Other publication FI 953578 960127  
Other publication CA 2154628 960127  
Other publication AU 2508695 960208.  
FEATURES  
source  
1. 744  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
CDS  
1. 744  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAA03092.1"  
/db\_xref="GI:2302074"  
translation="MRQKLTLLAFVCFALTPAPBIIQAOIIVTDSINHGDIYDE  
FWKDSGGSGTWIILNHLNFGVQVQNNVNIILFRGKKEFNETQTHQVGNMISYNGANFQ  
PNGNAYLCVVGWTVDPLEVEYIVDSWGNWRPPGATPKGTITVDGGTYDIYETLRVNP  
SILGIATFKOYWSVRRSKRTSGTISVSNHFRAMENLGMNWKMYEVALTVVEGYOSSGS  
ANYVNTLRINGNPLSTISNDESITLDKNN"  
sig\_peptide 1. 81  
mat\_peptide 82. 744  
/product="unnamed"  
ORIGIN  
Query Match 100.0%; Score 663; DB 6; Length 744;  
Best Local Similarity 100.0%; Pred. No. 2.1e-169; Mismatches 0; Indels 0; Gaps 0;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAATCGTCACCGACAAATTCATTCGCAACCAACGATGGCTATGATTATGAATTTGGAAA 60  
DB 82 CAATCGTCACCGACAAATTCATTCGCAACCAACGATGGCTATGATTATGAATTTGGAAA 141  
QY 61 GATAGCGGTGGCTCTGGGACAAATGATTCATCATGATGGGGTACGTTGAGTCCCAATGG 120  
DB 142 GATAGCGGTGGCTCTGGGACAAATGATTCATCATGATGGGGTACGTTGAGTCCCAATGG 201  
QY 121 AACATGTTAAACAATATTTTCGTAAGGTAAATAATTCATGAAACACAAACACAC 180  
DB 202 AACATGTTAAACAATATTTTCGTAAGGTAAATAATTCATGAAACACAAACACAC 261  
QY 181 CAACAAGTTGGTAACATGTCCTAATAAATAACGAGGCAACTTCACACCAAAATGGTAATGG 240  
DB 262 CAACAAGTTGGTAACATGTCCTAATAAATAACGAGGCAACTTCACACCAAAATGGTAATGG 321  
QY 241 TATTTATGCGTCTATGGTTGGACTGTTGACCTCTGTGCGATATATTTATGTCACAGT 300  
DB 322 TATTTATGCGTCTATGGTTGGACTGTTGACCTCTGTGCGATATATTTATGTCACAGT 381  
QY 301 TGGGGCAACTGGCGTCCACGAGGACCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 360  
DB 382 TGGGGCAACTGGCGTCCACGAGGACCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 441



```
QY 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAAGGGGATTGCCACA 420
DB 442 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAAGGGGATTGCCACA 501
QY 421 TTTAAACAATATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCACGATTTCTGTGACG 480
DB 502 TTTAAACAATATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCACGATTTCTGTGACG 561
QY 481 AACCACTTTAGACGCTGGGAAACTTAGGATGAATATGGGGAATATGATGAAGTCGCG 540
DB 562 AACCACTTTAGACGCTGGGAAACTTAGGATGAATATGGGGAATATGATGAAGTCGCG 621
QY 541 CTTACTGTAGAGCTTCAACTAGCGGAGTCTTAATGTATATAGCAATACACTAAGA 600
DB 622 CTTACTGTAGAGCTTCAACTAGCGGAGTCTTAATGTATATAGCAATACACTAAGA 681
QY 601 ATTAACGGTAAACCTCTCTCAACTATTAGTAATGACGAGAGCATAACTTTGGATAAAAAC 660
DB 682 ATTAACGGTAAACCTCTCTCAACTATTAGTAATGACGAGAGCATAACTTTGGATAAAAAC 741
QY 661 AAT 663
DB 742 AAT 744

RESULT 7
AR193051
LOCUS AR193051 744 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 4 from patent US 6346407.
ACCESSION AR193051
VERSION AR193051.1 GI:20239016
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 744)
AUTHORS De Buyl,E., Lahaye,A., Ledoux,P. and Detroz,R.
TITLE Xylanase, microorganisms producing it, DNA molecules, methods for preparing this xylanase and uses of the latter
JOURNAL Patent: US 6346407-A 4 12-FEB-2002;
FEATURES
source
1..744
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 663; DB 6; Length 744;
Best Local Similarity 100.0%; Pred. No. 2.1e-169;
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAATCGTCACCGACAATTCATTTGGCAACACGATGGCTATGATTATGAAATTTGGAAA 60
DB 82 CAAATCGTCACCGACAATTCATTTGGCAACACGATGGCTATGATTATGAAATTTGGAAA 141
QY 61 GATAGCGGTGGCTCTGGGACAATGATTTCTCAATCATGGCGGTACGTTTCAGTGCCTAATGG 120
DB 142 GATAGCGGTGGCTCTGGGACAATGATTTCTCAATCATGGCGGTACGTTTCAGTGCCTAATGG 201
QY 121 AACCAATGTTAAACAATATTATTCGTTAAAGGTAAAGGATTTCAATGAAGAACACAAACAC 180
DB 202 AACCAATGTTAAACAATATTATTCGTTAAAGGTAAAGGATTTCAATGAAGAACACAAACAC 261
QY 181 CAACAAGTTGGTAAACATGTCATAAATAGCGAGCAACTTCCAAACCAATGTTAATGCG 240
DB 262 CAACAAGTTGGTAAACATGTCATAAATAGCGAGCAACTTCCAAACCAATGTTAATGCG 321
QY 241 TATTTATGCGTCTATGTTGACCTCTTGTGCGAATATTTATTTGTCGACAGT 300
DB 322 TATTTATGCGTCTATGTTGACCTCTTGTGCGAATATTTATTTGTCGACAGT 381
QY 301 TGGGGCAACTGGCGTCCACGAGGAGCAACGCTTAAGGGGACCACTCACTGTTGATGGAGGA 360
```

```
DB 382 TGGGGCAACTGGCGTCCACGAGGAGCAACGCCCTAAAGGGGACCATCACTGTTGATGGAGGA 441
QY 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAAGGGGATTGCCACA 420
DB 442 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAAGGGGATTGCCACA 501
QY 421 TTTAAACAATATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCACGATTTCTGTGACG 480
DB 502 TTTAAACAATATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCACGATTTCTGTGACG 561
QY 481 AACCACTTTAGACGCTGGGAAACTTAGGATGAATATGGGGAATATGATGAAGTCGCG 540
DB 562 AACCACTTTAGACGCTGGGAAACTTAGGATGAATATGGGGAATATGATGAAGTCGCG 621
QY 541 CTTACTGTAGAGCTTCAACTAGCGGAGTCTTAATGTATATAGCAATACACTAAGA 600
DB 622 CTTACTGTAGAGCTTCAACTAGCGGAGTCTTAATGTATATAGCAATACACTAAGA 681
QY 601 ATTAACGGTAAACCTCTCTCAACTATTAGTAATGACGAGAGCATAACTTTGGATAAAAAC 660
DB 682 ATTAACGGTAAACCTCTCTCAACTATTAGTAATGACGAGAGCATAACTTTGGATAAAAAC 741
QY 661 AAT 663
DB 742 AAT 744

RESULT 8
AR193052
LOCUS AR193052 744 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 5 from patent US 6346407.
ACCESSION AR193052
VERSION AR193052.1 GI:20239017
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 744)
AUTHORS De Buyl,E., Lahaye,A., Ledoux,P. and Detroz,R.
TITLE Xylanase, microorganisms producing it, DNA molecules, methods for preparing this xylanase and uses of the latter
JOURNAL Patent: US 6346407-A 5 12-FEB-2002;
FEATURES
source
1..744
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 663; DB 6; Length 744;
Best Local Similarity 100.0%; Pred. No. 2.1e-169;
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAATCGTCACCGACAATTCATTTGGCAACACGATGGCTATGATTATGAAATTTGGAAA 60
DB 82 CAAATCGTCACCGACAATTCATTTGGCAACACGATGGCTATGATTATGAAATTTGGAAA 141
QY 61 GATAGCGGTGGCTCTGGGACAATGATTTCTCAATCATGGCGGTACGTTTCAGTGCCTAATGG 120
DB 142 GATAGCGGTGGCTCTGGGACAATGATTTCTCAATCATGGCGGTACGTTTCAGTGCCTAATGG 201
QY 121 AACCAATGTTAAACAATATTATTCGTTAAAGGTAAAGGATTTCAATGAAGAACACAAACAC 180
DB 202 AACCAATGTTAAACAATATTATTCGTTAAAGGTAAAGGATTTCAATGAAGAACACAAACAC 261
QY 181 CAACAAGTTGGTAAACATGTCATAAATAGCGAGCAACTTCCAAACCAATGTTAATGCG 240
DB 262 CAACAAGTTGGTAAACATGTCATAAATAGCGAGCAACTTCCAAACCAATGTTAATGCG 321
QY 241 TATTTATGCGTCTATGTTGACCTCTTGTGCGAATATTTATTTGTCGACAGT 300
DB 322 TATTTATGCGTCTATGTTGACCTCTTGTGCGAATATTTATTTGTCGACAGT 381
QY 301 TGGGGCAACTGGCGTCCACGAGGAGCAACGCTTAAGGGGACCACTCACTGTTGATGGAGGA 360
```

Db 382 TGGGGCAATGGCGTCCACCGAGGACACCGCTTAAGGGACCATCACTGTTGATGGAGGA 441  
Qy 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTGCCACA 420  
Db 442 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTGCCACA 501  
Qy 421 TTTAAACAATATGGAGTGTTCGAAGATCGAAACGACGAGTGGACGATTTCTGTGAGC 480  
Db 502 TTTAAACAATATGGAGTGTTCGAAGATCGAAACGACGAGTGGACGATTTCTGTGAGC 561  
Qy 481 AACCACTTTAGAGCGTGGGAAACCTTAGGATGAATATGGGAAATGTATGAAGTGGCG 540  
Db 562 AACCACTTTAGAGCGTGGGAAACCTTAGGATGAATATGGGAAATGTATGAAGTGGCG 621  
Qy 541 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA 600  
Db 622 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA 681  
Qy 601 ATTACGGTAAACCTCTCTCAACTATTAGTAAATGACGAGACATAACTTTGGATAAAAC 660  
Db 682 ATTACGGTAAACCTCTCTCAACTATTAGTAAATGACGAGACATAACTTTGGATAAAAC 741  
Qy 661 AAT 663  
Db 742 AAT 744

RESULT 9  
LOCUS A48231 1513 bp DNA linear PAT 07-MAR-1997  
DEFINITION Sequence 10 from Patent EP0698667.  
ACCESSION A48231  
VERSION A48231.1 GI:2302078  
KEYWORDS  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 1513)  
AUTHORS De,B.E., Lahaye,A., Ledoux,P. and Detroz,R.  
TITLE Xylanase, microorganisms for its production, DNA molecules, process of preparation and use thereof  
JOURNAL Patent: EP 0698667-A 10 28-FEB-1996;  
COMMENT SOLVAY (BE)  
Other publication BE 1008751 960702  
Other publication BE 1008570 960604  
Other publication BR 9503454 960305  
Other publication JP 8092284 960409  
Other publication FI 953578 960127  
Other publication CA 2154628 960127  
Other publication AU 2508695 960208.

FEATURES  
source  
1. .1513  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

ORIGIN  
Query Match 100.0%; Score 663; DB 6; Length 1513;  
Best Local Similarity 100.0%; Pred. No. 2e-169;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAAATCGTCACGACAAATTCATTGGCAACCGAGTGGCTATGATTAATGAATTTGGAAA 60  
Db 701 CAAATCGTCACGACAAATTCATTGGCAACCGAGTGGCTATGATTAATGAATTTGGAAA 760  
Qy 61 GATAGCGGTGGCTCTGGGACAATGATTTCAATCATGTCGGGTACGTTTCAGTCCCAATGG 120  
Db 761 GATAGCGGTGGCTCTGGGACAATGATTTCAATCATGTCGGGTACGTTTCAGTCCCAATGG 820  
Qy 121 AACAAATGTTAACACATATTTATTCCTGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGG 180  
Db 821 AACAAATGTTAACACATATTTATTCCTGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGG 880

Qy 181 CACAAAGTTGGTAACATGTCTTAATACTAGGAGCCAACTTCCAAACCAATGGTAATGCG 240  
Db 881 CACAAAGTTGGTAACATGTCTTAATACTAGGAGCCAACTTCCAAACCAATGGTAATGCG 940  
Qy 241 TATTTATGGCTCTATGGTTGGACTGTGTGACCTCTTGTGCGAATATATATTTGTCGACGT 300  
Db 941 TATTTATGGCTCTATGGTTGGACTGTGTGACCTCTTGTGCGAATATATATTTGTCGACGT 1000  
Qy 301 TGGGGCAACTGGGCTCCACAGAGCAACGCCCTTAAGGGGACCATCACTGTTGATGGAGGA 360  
Db 1001 TGGGGCAACTGGGCTCCACAGAGCAACGCCCTTAAGGGGACCATCACTGTTGATGGAGGA 1060  
Qy 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTGCCACA 420  
Db 1061 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTGCCACA 1120  
Qy 421 TTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGGATTTCTGTGAGC 480  
Db 1121 TTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGGATTTCTGTGAGC 1180  
Qy 481 AACCACTTTAGAGCGTGGGAAACCTTAGGATGAATATGGGAAATGTATGAAGTGGCG 540  
Db 1181 AACCACTTTAGAGCGTGGGAAACCTTAGGATGAATATGGGAAATGTATGAAGTGGCG 1240  
Qy 541 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA 600  
Db 1241 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA 1300  
Qy 601 ATTACGGTAAACCTCTCTCAACTATTAGTAAATGACGAGACATAACTTTGGATAAAAC 660  
Db 1301 ATTACGGTAAACCTCTCTCAACTATTAGTAAATGACGAGACATAACTTTGGATAAAAC 1360  
Qy 661 AAT 663  
Db 1361 AAT 1363

RESULT 10  
LOCUS A48232 1513 bp DNA linear PAT 07-MAR-1997  
DEFINITION Sequence 11 from Patent EP0698667.  
ACCESSION A48232  
VERSION A48232.1 GI:2302079  
KEYWORDS  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 1513)  
AUTHORS De,B.E., Lahaye,A., Ledoux,P. and Detroz,R.  
TITLE Xylanase, microorganisms for its production, DNA molecules, process of preparation and use thereof  
JOURNAL Patent: EP 0698667-A 11 28-FEB-1996;  
COMMENT SOLVAY (BE)  
Other publication BE 1008751 960702  
Other publication BE 1008570 960604  
Other publication BR 9503454 960305  
Other publication JP 8092284 960409  
Other publication FI 953578 960127  
Other publication CA 2154628 960127  
Other publication AU 2508695 960208.

FEATURES  
source  
1. .1513  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
620- .1366  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAA03094.1"  
/db\_xref="GI:2302080"  
/translation="MRQKLTLLAFLVCFALTLPAEIIQAIQIVTDSIGNHDGYDE  
PWKDSGSGSTMILNHEGTFSAQWNNVNNILFRKGKFNFTQHQVGNMINSYNGANFQ



1 (bases 1 to 744)  
Dalboege,H., Diderichsen,B., Sandal,T. and Kauppinen,S.  
METHOD OF PROVIDING NOVEL DNA SEQUENCES  
Patent: WO 9743409-A 1 20-NOV-1997;  
JOURNAL

NOVONORDISK AS (DK)  
 FEATURES  
 Location/Qualifiers  
 1. .744  
 /organism="unidenti"

```
/mol_type="unassigned DNA"
/strain="BACILLUS SP. AC13"
/db_xref="taxon:32644"
```

```

/ww_xref= taxon:3264
1. .>744
/note="unnamed protein product"

```

```
/codon_start=1
/protein_id="CAB42305.1"
```

```

/db_xref="GI:4756811"
/translation="MRQKLTFILAFVCFALTLP AEITQAIQIVTDSIGNHDGYDYE

```

[illegible]

ANVYSNTRIRNGNPLSTISNDKSI TLDKNN"

95.7%; Score 634.2; DB 6; Length 744;

Local Similarity	97.3%	Pred. No. 1.4e-161;	
Matches 645;	Conservative	0;	Mismatches 18;
		Indels	0;
		Gaps	0;

1 CAAATCGTCACCGACAAATCCATTGGCAACCAACGATGGCTATGATTATGAATTTGGAAA 60

82 CAAATCGTCACCGACAATTCCATTGGCAACCAACGATGGCTATGATTATGAATTTGGAA 141

61 GATAGCGGTGGCTCTGGGACAATGATTCTCAATCATGGCGGTACGTTCAGTGCCCAATGG 120

142 GATAGCGGTGCTCTGGACAAATGATTCTCAATCATGGCGGTACGTTAGTGCCCAATGG 201

121 AACAAATGTTAAACAACATATTATTCCTGTAAGGTAAAAATTCATAGTAAACACAAACACAC 180

202 AACAAATTAAACAACATATTATTCGTAAAGGTAAAAAATTCATGTGAAACACACAAACACAC 261

[illegible]

262 CACAAAGTGGGTAACAATGTCCCAATTACTTAGGCCCTTTTGTCGCACAGT' 300

241 T'AATTTA TGCGCTC'TA TGSTNGGA CGNTTGA CCTCTTGT GCAATA TTA TAT TTGTCGACAGT'

241 TATTATCGGCTCATGTTGGACCTCTTGTGGAATATTATATTCTCATAGT 381

301 TGGGGCAACTGGCGTCCACCAGGCAACGCTAAGGGGACCATCACTGTTGATGAGGA 360

382 TGGGGCAACTGGCGTGCCACAGGGGCAAGCCTAAGGGAAACCATCACTGTTGATGGAGGA 441

361 ACATATGATATCTACGAGACTTTAGAGTCAATCAACCTCCATTAGGGGATTGCCACA 420

442 ACATATGATATCTATGAAACTCTTAGAGTCAATCAGCCCTTCATTAGGGGATTGCCACA 501

421 TTTTAAACAATATTGGAGTGTTCGAAGATCGAATCGCACGAGTGGCACAATTTCTGTTCAGC 480

502 TTTAAACAATATTGGAGTCTCCGAAGATCGAACGCGCAGTAGGCGACAATTTCTGTGCAGC 561

481 AACCACTTTAGACGCTGGGAAAATCTTAGGGCATGATAATGGGGAAAAATGTAATGAAGATGCGG 54

562 AACACACTTTAGACGCGTGGGAATACTTAGGGGATGACATGCCGCTTGATGTGTCGCTCCG  
541 GTTATCCTCTAGAACGGCTATCAAAGTAGCGGAAGTGTATGTATATAGCAATACACTTAAGA 600

341 CTTACTGTAGAGGCTATCAAAGTAGCGGAAGTGCTAATGTATATAGCAATACACTAAGA 681  
622 CTTACTGTAGAGGCTATCAAAGTAGCGGAAGTGCTAATGTATATAGCAATACACTAAGA 681

601 ATTAACGGTAACCTCTCTCAACTATTAGTAATGACGAGAGCATAACTTTGGATAAAAC 660

682 ATTAACGGTAACCTCTCTCAACTATTAGTAATGA CAAGAGCATAACTTAGATAAAAC 741

661 AAT 663

```
Db 742 AAT 744

RESULT 14
LOCUS AR163110
DEFINITION Sequence 1 from patent US 6270968.
ACCESSION AR163110
VERSION AR163110.1 GI:16233600
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 744)
AUTHORS Dalb.o slashed.ge,H., Sandal,T., Kauppinen,M.Sakari. and
Diderichsen,Belashedge.
TITLE Method of providing a hybrid polypeptide exhibiting an activity of
interest
JOURNAL Patent: US 6270968-A 1 07-AUG-2001;
FEATURES
source
1..744
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 95.7%; Score 634.2; DB 6; Length 744;
Best Local Similarity 97.3%; Pred. No. 1.4e-161;
Matches 645; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 CAAATCGTCACCGACAATTCATTTGGCAACACGATGGCTATGATTATGAATTTGGAAA 60
Db 82 CAAATCGTCACCGACAATTCATTTGGCAACACGATGGCTATGATTATGAATTTGGAAA 141
QY 61 GATAGCGGTGGCTCTGGGCAATGATCTCAATCATATGGCGGTACGTTCAAGTCCCAATGG 120
Db 142 GATAGCGGTGGCTCTGGGCAATGATCTCAATCATATGGCGGTACGTTCAAGTCCCAATGG 201
QY 121 AACATGTTAAACAATATATTTCCGTAAGGTAAAGGTAAAGATTCATTAAGAAACACACAC 180
Db 202 AACATGTTAAACAATATATTTCCGTAAGGTAAAGGTAAAGATTCATTAAGAAACACACAC 261
QY 181 CAACAAGTTGGTAAACATGTCATATAACTACGAGGACCAACTTCCAAACAAATGGTAAATGG 240
Db 262 CAACAAGTTGGTAAACATGTCATATAACTACGAGGACCAACTTCCAAACAAACGGAATGG 321
QY 241 TATTTATGGCTATATGGTGGAGCTGTGACCCCTCTTGTGCAATATATATTTGTCGACAGT 300
Db 322 TATTTATGGCTATATGGTGGAGCTGTGACCCCTCTTGTGCAATATATATTTGTCGATAGT 381
QY 301 TGGGGCAACTGGCGTCCACGAGCAACGCCCTAAAGGGACCACTCACTGTTGATGGAGGA 360
Db 382 TGGGGCAACTGGCGTCCACGAGGCAACGCCCTAAAGGGACCACTCACTGTTGATGGAGGA 441
QY 361 ACATATGATATCTACAGAGCTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTCGCACA 420
Db 442 ACATATGATATCTATGAAGCTCTTAGAGTCAATCAGCCCTCCATTAAGGGGATTCGCACA 501
QY 421 TTTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGCAAGTGGCGACGATTTCTGTACG 480
Db 502 TTTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGCAAGTGGCGACCAATTTCTGTACG 561
QY 481 AACCACTTTAGAGCGTGGGAAACTTAGGATGAATATGGGAAATATGATGAAGTCGG 540
Db 562 AACCACTTTAGAGCGTGGGAAACTTAGGATGAATATGGGAAATATGATGAAGTCGG 621
QY 541 CTTTACTGTAGAAGCTATCAAGTGTAGCGGAAGTGTAAATGATATATAGCAATACACTAAGA 600
Db 622 CTTTACTGTAGAAGCTATCAAGTGTAGCGGAAGTGTAAATGATATATAGCAATACACTAGA 681
QY 601 ATTTAACGGTAAACCTCTCTCAACTATTTAGTAATGACGAGAGCATTAATTTGGATATAAAC 660
Db 682 ATTTAACGGTAAACCTCTCTCAACTATTTAGTAATGACGAGAGCATTAATTTAGATATAAAC 741

QY 661 AAT 663
Db 742 AAT 744

RESULT 15
LOCUS A45313
DEFINITION Sequence 18 from Patent WO9518219.
ACCESSION A45313
VERSION A45313.1 GI:2299796
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 744)
AUTHORS Van,S.P., Williams,D.P., Iverson,S., Farrell,R.L., Herbes,W.T.,
Van,D.K., Herwijer,M.A., Van,B.R., Quax,W.J., Goedegebuur,F. and
Jones,B.E.
TITLE ALKALI-TOLERANT XYLANASES
JOURNAL Patent: WO 9518219-A 18 06-JUL-1995;
COMMENT GIST BROCADES NV (NL)
Other publication JP 8507221T 960806
Other publication BR 9405934 951226
Other publication NO 953312 951019
Other publication FI 953920 950821
Other publication AU 1415095 950717.
FEATURES
Location/Qualifiers
1..744
/organism="unidentified"
/mol_type="unassigned DNA"
/strain="1-43-3"
/isolate="CBS672.93"
/db_xref="taxon:32644"
1..744
/codon_start=1
/product="XYLANASE"
/protein_id="CAA02821.1"
/db_xref="GI:2299797"
/translation="MSOKKLTLINLFLPALTPARISOAQIVTNSIATRGSDYEF
WKDSGSGTMILNHGSGTSAQWNNVNLPRKGFNETQHQOQVGNMSINYGANFQP
NGNAYLVGVMTVDLVEYIIVDSMGNWRPPGATPKGTTVDGGTVDIYETLARVQPS
IKGIATFKYMSVRSKTSKTSISVSNHFRWENLGMNMGMKMEVALTVGEGYQSSGA
NVYSNTRLINGNPLSTISNNEITLDKNN"

ORIGIN
Query Match 94.2%; Score 624.6; DB 6; Length 744;
Best Local Similarity 96.4%; Pred. No. 5.8e-159;
Matches 639; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 CAAATCGTCACCGACAATTCATTTGGCAACACGATGGCTATGATTATGAATTTGGAAA 60
Db 79 CAAATCGTCACCGACAATTCATTTGGCAACACGATGGCTATGATTATGAATTTGGAAA 138
QY 61 GATAGCGGTGGCTCTGGGCAATGATCTCAATCATATGGCGGTACGTTCAAGTCCCAATGG 120
Db 139 GATAGCGGTGGCTCTGGGCAATGATCTCAATCATATGGCGGTACGTTCAAGTCCCAATGG 198
QY 121 AACATGTTAAACAATATATTTCCGTAAGGTAAAGGTAAAGATTCATTAAGAAACACACAC 180
Db 199 AATAATGTTAAACAATATATTTCCGTAAGGTAAAGGTAAAGATTCATTAAGAAACACACAC 258
QY 181 CAACAAGTTGGTAAACATGTCATATAACTACGAGGACCAACTTCCAAACAAATGGTAAATGG 240
Db 259 CAACAAGTTGGTAAACATGTCATATAACTACGAGGACCAACTTCCAAACAAACGGAATGG 318
QY 241 TATTTATGGCTATATGGTGGAGCTGTGACCCCTCTTGTGCAATATATATTTGTCGACAGT 300
Db 319 TATTTATGGCTATATGGTGGAGCTGTGACCCCTCTTGTGCAATATATATTTGTCGACAGT 378
QY 301 TGGGGCAACTGGCGTCCACGAGGACCAACGCCCTAAAGGGACCACTCACTGTTGATGGAGGA 360
Db 379 TGGGGCAACTGGCGTCCACGAGGACCAACGCCCTAAAGGGACCACTCACTGTTGATGGAGGA 438
```

Qy	361	ACATATGATCTATAGAGACTCTTAGAGTCAATCAACCCCTCCATTAAAGGGATTGCCACA	420
Db	439	ACATATGATATCTATGAAACTCTTAGAGTCAATCAGCCCTCCATTAAAGGGATTGCCACA	498
Qy	421	TTTAAACAATATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGATTCTGTGAGC	480
Db	499	TTTAAACAATATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGATTCTGTGAGC	558
Qy	481	AACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATGGGAAATGTATGAAAGTCGCG	540
Db	559	AACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATGGGAAATGTATGAAAGTCGCG	618
Qy	541	CTTACTGTAGAAAGGCTATCAAAGTAGCGGAAGTCTAATGTATATAGCAATACACTAAGA	600
Db	619	CTTACTGTAGAAAGGCTATCAAAGTAGCGGAAGTCTAATGTATATAGCAATACACTAAGA	678
Qy	601	ATTACGGTAAACCTCTCTCAACTATTAGTAAATGACGAGAGCATAACTTTGGATAAAAAAC	660
Db	679	ATTACGGTAAACCTCTCTCAACTATTAGTAAATGACGAGAGCATAACTTTGGATAAAAAAC	738
Qy	661	AAT 663	
Db	739	AAT 741	

Search completed: November 10, 2004, 21:26:27  
 Job time : 3210.21 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 16:24:32 ; Search time 801.957 Seconds  
(without alignments)  
8406.125 Million cell updates/sec

Title: US-09-909-207-26  
Perfect score: 185  
Sequence: 1 TCATGTAACCTCCCTTGATC.....AAAGAGGAGGATGGA 185

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hcc1:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	37.6	20.3	1281	9	AG346331 Mus muscu
2	37.4	20.2	865	9	AG388392 Mus muscu
3	37.2	20.1	628	1	AL867917 AL867917
4	37.2	20.1	659	1	AL846715 AL846715
5	37.2	20.1	663	5	BQ527682 NISC.no23
6	37.2	20.1	880	7	CR584662 CR584662
C 7	37.2	20.1	955	5	EX772034 BX772034
8	36	19.5	909	8	AZ549992 ENTGS63TF
9	35.8	19.4	714	9	CE215054 tigr-gss-
C 10	35.8	19.4	1101	9	CNS017L2 Drosophil
11	35.6	19.2	352	7	CR449406 CR449406
C 12	35.6	19.2	833	8	B2224970 CH230-351
C 13	35.6	19.2	874	7	CR579359 CR579359
C 14	35.6	19.2	960	1	AL901596 AL901596
15	35.4	19.1	951	4	BG164591 BG164591
C 16	35.2	19.0	732	9	AG343254 Mus muscu
17	35	18.9	444	7	CN450127 GUO cdNA
18	35	18.9	686	7	CN450126 GUO cdNA
19	35	18.9	911	8	CC386284 PUHTN23TB
20	34.8	18.8	500	8	AQ154047 HS_2236_B
21	34.6	18.7	485	2	BF702386 MI-P-H3-a
C 22	34.6	18.7	523	4	BJ644137 BJ644137
C 23	34.6	18.7	703	4	BJ641803 BJ641803
24	34.6	18.7	867	4	BG539277 BG539277

C 25	34.6	18.7	921	9	CL264923
C 26	34.4	18.6	203	8	AZ037080
C 27	34.2	18.5	529	7	CN055278 Salamande
C 28	34	18.4	492	9	CE103446 tigr-gss-
C 29	34	18.4	593	2	AW645414 cm52f09.w
C 30	34	18.4	636	9	AL982160 Danio rer
C 31	34	18.4	796	9	CL283976 ZMMBB062
C 32	34	18.4	1029	5	BX424910 BX424910
C 33	33.8	18.3	619	5	BW308307 BW308307
C 34	33.8	18.3	679	6	CF180381 815477 MA
C 35	33.8	18.3	681	9	CL551099 OB_Ba009
C 36	33.8	18.3	697	8	BH965893 odh5h05.
C 37	33.8	18.3	721	8	BH965366 ood95b08.
C 38	33.8	18.3	724	9	CL509951 SAIL_822
C 39	33.8	18.3	737	5	BW183846 BW183846
C 40	33.6	18.2	536	9	CE133665 tigr-gss-
C 41	33.6	18.2	570	8	AZ908959 RPI-24-1
C 42	33.6	18.2	609	9	CE263180 tigr-gss-
C 43	33.6	18.2	621	9	CE202395 tigr-gss-
C 44	33.6	18.2	700	8	AZ909070 RPI-24-1
C 45	33.6	18.2	733	9	AG380466 Mus muscu

ALIGNMENTS

RESULT 1  
AG346331/c  
LOCUS  
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-141H11.T7, genomic survey  
sequence.  
ACCESSION AG346331 GI:47919641  
VERSION AG346331.1  
KEYWORDS GSS.  
SOURCE Mus musculus molossinus  
ORGANISM Mus musculus molossinus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1  
AUTHORS Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.  
TITLE BAC end Sequences of Library MSMg01  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1281)  
AUTHORS Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.  
TITLE Direct Submission  
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/,  
Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC  
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).  
Tsukuba Institute, Bio Resource Center, Research (RIKEN) 3-1-1  
Koyada, Tsukuba, 305-0074 Japan  
phone: 81-298-36-9189, fax: 81-298-36-9199  
e-mail: abe@rtc.riken.jp  
PRIMERS  
Sequencing : T7  
LIBRARY  
Vector : pBACe3.6  
R.Site 1 : EORI  
R.Site 2 : EORI.  
FEATURES  
Location/Qualifiers  
1. : 1281  
/organism="Mus musculus molossinus"  
/mol\_type="genomic DNA"  
/sub\_species="molossinus"  
/db\_xref="taxon:57486"  
/clone="MSMg01-141H11.T7"  
/sex="male"  
/tissue\_type="mixture of kidney and spleen"  
/clone\_lib="MSMg01 Mouse Male BAC Library"



```

ORIGIN
Query Match      20.3%; Score 37.6; DB 9; Length 1281;
Best Local Similarity 52.7%; Pred. No. 2.1;
Matches 79; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 36 AAAGGATTATACACAAACAGAGACATCCATGCCGGTTAAAGCAGTATCGTTCCATCT 95
DB 1225 AAAGAAGGTGAAGATCAAGAGAAATACCATGCCAAGAGTAAGAGATAAATCTTTTAA 1166
QY 96 AACAGAGAAAGNCTGCATCAAGAGAGGTGATGGCTTTTTCATCTTAGGGATGACAGACA 155
DB 1165 AAGATAGATAAATAAGAGAAATAGATGATATAATGNTATTGAAGTTGAGAGACAAAAA 1106
QY 156 ATACGGATCAAAAGAGAGAGAGGTGGGAAA 185
DB 1105 ATAGGATGAGAGAGAGAAAGTAATGAAA 1076

RESULT 2
AG388392      865 bp DNA linear GSS 03-JUN-2004
LOCUS Mus musculus molossinus DNA, clone:MSMg01-203L20.TJ, genomic survey
DEFINITION
ACCESSION AG388392.1 GI:47999597
VERSION AG388392
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
REFERENCE 1 Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
AUTHORS BAC end Sequences of Library MSMg01
TITLE Unpublished
JOURNAL 2 (bases 1 to 865)
REFERENCE Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
AUTHORS Direct Submission
TITLE Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
JOURNAL and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY : pBACe3.6
Vector :
R.Site 1 : EcoRI
R.Site 2 : EcoRI
FEATURES Location/Qualifiers
source 1..865
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-203L20.TJ"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
Query Match      20.2%; Score 37.4; DB 9; Length 865;
Best Local Similarity 50.3%; Pred. No. 2.2;
Matches 83; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 21 TATTTTCATTTGTATCAAGGATTTTATACAAACAGAGACATCCATGCCGGTTAAAGC 80

```

```

DB 8 TATTTTAAAAAGGAAAAAAGGAAAAAGGAAAAAGGAAAAAGGAAAAAGGAAAAAGG 67
QY 81 AGTATCGTTCCATCTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 140
DB 68 CGGNNNGNGTTCATCAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 127
QY 141 AGGGATGACAGACATACGATGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 185
DB 128 AAAGGGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 172

RESULT 3
AL867917      628 bp mRNA linear EST 03-DEC-2003
LOCUS AL867917 XGC-egg Xenopus tropicalis cDNA clone TEGg105m20 5', mRNA
DEFINITION
ACCESSION AL867917
VERSION AL867917
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 628)
AUTHORS Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
TITLE Sanger Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL Unpublished (2003)
COMMENT On Sep 15, 2002 this sequence version replaced gi:22888182.
Contact: Taylor R
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TEGg105m20.plkSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from Sug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site1: EcoRI; Site 2: NotI
Host: Escherichia coli Xli-blue.
FEATURES Location/Qualifiers
source 1..628
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TEGg105m20"
/dev_stage="egg"
/lab_host="Escherichia coli Xli-blue"
/clone_lib="XGC-egg"
/notes="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from Sug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end"

ORIGIN
Query Match      20.1%; Score 37.2; DB 1; Length 628;
Best Local Similarity 58.9%; Pred. No. 2.4;
Matches 63; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 16 TGATCTATTTTCATTTGTATCAAGGATTTTATACAAACAGAGACATCCATGCCGGTT 75
DB 500 TGTACTAGATAAAATATCCCAAGAAATTTATAAGAACAAAGAACTATTATGAGAAAT 559
QY 76 AAAGCAGTATCGTTCCATCTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 122
DB 560 GAAGGAGTATGCCCTGCTACACTTTTGTGCTGTCATCTGAGGAAG 606

RESULT 4

```

AL846715 659 bp mRNA linear EST 20-NOV-2003  
 LOCUS AL846715 XGC-egg Xenopus tropicalis cDNA clone TEGG049119 5', mRNA  
 DEFINITION sequence.  
 ACCESSION AL846715  
 VERSION AL846715  
 KEYWORDS EST.  
 SOURCE Xenopus tropicalis (western clawed frog)  
 ORGANISM Xenopus tropicalis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 Xenopodinae; Xenopus; Silurana.  
 1 (bases 1 to 659)  
 Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.  
 Sanger Xenopus tropicalis EST project 2001 (11\_2003)  
 TITLE Unpublished (2003)  
 JOURNAL  
 COMMENT On Sep 15, 2002 this sequence version replaced gi:22866980.  
 Contact: Taylor R  
 Sanger Institute  
 Hinxton, Cambridgeshire, CB10 1SA, UK  
 Email: trop@sanger.ac.uk  
 Sanger Xenopus tropicalis EST project 2001  
 TROPICALIS SEQUENCE ID: TEGG049119.plksp6w  
 Sequencing primer: SP6  
 This sequence is from a Xenopus Gene Collection (XGC) library  
 constructed by Aaron M. Zorn.  
 cDNA was oligo dt primed from Sug of poly A+ RNA from egg.  
 EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the  
 5' end and NotI at the 3' end.  
 Vector: pCS107; Site 1: EcoRI; Site 2: NotI  
 Host: Escherichia coli XL1-blue.  
 Location/Qualifiers  
 1..659  
 /organism="Xenopus tropicalis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8364"  
 /clone="TEGG049119"  
 /dev\_stage="egg"  
 /lab\_hosts="Escherichia coli XL1-blue"  
 /clone\_lib="XGC-egg"  
 /note="vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA  
 was oligo dt primed from Sug of poly A+ RNA from egg.  
 EcoRI-NotI cut cDNA was then ligated into pCS107 with  
 EcoRI at the 5' end and NotI at the 3' end"

Query Match 20.1%; Score 37.2; DB 1; Length 659;  
 Best Local Similarity 58.9%; Pred. No. 2.5;  
 Matches 63; Conservative 0; Mismatches 44; Indels 0; Gaps 0;  
 Qy 16 TGATCTATTTTCATTTGATCAAGGATTTATACAAAGAGAGACATCCATCCCGGTT 75  
 Db 421 TGTACTAGATAAAATATCCCAAGATTTTATAAGAACAAAGACTATTATGAGGAAT 480  
 Qy 76 AAGCAGTATGTTCCATCTAACAGAGAGAGGCTGCTGCACTGAGGAGG 122  
 Db 481 GAAGGAGTATGGCCCTGCTACACTTTGTGTCTGCACTGAGGAGG 527

## ORIGIN

Query Match 20.1%; Score 37.2; DB 1; Length 659;  
 Best Local Similarity 58.9%; Pred. No. 2.5;  
 Matches 63; Conservative 0; Mismatches 44; Indels 0; Gaps 0;  
 Qy 16 TGATCTATTTTCATTTGATCAAGGATTTATACAAAGAGAGACATCCATCCCGGTT 75  
 Db 421 TGTACTAGATAAAATATCCCAAGATTTTATAAGAACAAAGACTATTATGAGGAAT 480  
 Qy 76 AAGCAGTATGTTCCATCTAACAGAGAGAGGCTGCTGCACTGAGGAGG 122  
 Db 481 GAAGGAGTATGGCCCTGCTACACTTTGTGTCTGCACTGAGGAGG 527

RESULT 5  
 BQ527682  
 LOCUS BQ527682 663 bp mRNA linear EST 10-JUN-2002  
 DEFINITION NISC no23g06.y1 NICHG\_XGC\_Emb8 Xenopus tropicalis cDNA clone  
 IMAGE:5381962 5', mRNA sequence.  
 ACCESSION BQ527682  
 VERSION BQ527682  
 KEYWORDS EST.  
 SOURCE Xenopus tropicalis (western clawed frog)  
 ORGANISM Xenopus tropicalis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 Xenopodinae; Xenopus; Silurana.  
 1 (bases 1 to 663)

AUTHORS  
TITLE  
JOURNAL  
COMMENT

NIH-XGC [http://image.llnl.gov/image/html/xenopuslib\\_info.shtml](http://image.llnl.gov/image/html/xenopuslib_info.shtml).  
 National Institute of Child Health and Human Development, National  
 Cancer Institute, Xenopus Gene Collection  
 Unpublished (2002)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 cDNA Library Preparation:  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC)  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM11972 row: N column: 11  
 Seq primer: M13RP1 reverse primer (ABI).  
 Location/Qualifiers  
 1..663  
 /organism="Xenopus tropicalis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8364"  
 /clone="IMAGE:5381962"  
 /tissue\_type="tadpole"  
 /dev\_stage="embryo, stages 40-45"  
 /lab\_hosts="DH10B (phage-resistant)"  
 /clone\_lib="NICHG\_XGC\_Emb8"  
 /note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV;  
 Cloned unidirectionally. Primer: Oligo dT. Average insert  
 size 2.1 kb. Constructed by Invitrogen. Note: This is a  
 Xenopus Gene Collection (XGC) library."

FEATURES  
source

Query Match 20.1%; Score 37.2; DB 5; Length 663;  
 Best Local Similarity 58.9%; Pred. No. 2.5;  
 Matches 63; Conservative 0; Mismatches 44; Indels 0; Gaps 0;  
 Qy 16 TGATCTATTTTCATTTGATCAAGGATTTATACAAAGAGAGACATCCATCCCGGTT 75  
 Db 345 TGTACTAGATAAAATATCCCAAGATTTTATAAGAACAAAGACTATTATGAGGAAT 404  
 Qy 76 AAGCAGTATGTTCCATCTAACAGAGAGGCTGCTGCACTGAGGAGG 122  
 Db 405 GAAGGAGTATGGCCCTGCTACACTTTGTGTCTGCACTGAGGAGG 451

## ORIGIN

Query Match 20.1%; Score 37.2; DB 5; Length 663;  
 Best Local Similarity 58.9%; Pred. No. 2.5;  
 Matches 63; Conservative 0; Mismatches 44; Indels 0; Gaps 0;  
 Qy 16 TGATCTATTTTCATTTGATCAAGGATTTATACAAAGAGAGACATCCATCCCGGTT 75  
 Db 345 TGTACTAGATAAAATATCCCAAGATTTTATAAGAACAAAGACTATTATGAGGAAT 404  
 Qy 76 AAGCAGTATGTTCCATCTAACAGAGAGGCTGCTGCACTGAGGAGG 122  
 Db 405 GAAGGAGTATGGCCCTGCTACACTTTGTGTCTGCACTGAGGAGG 451

## RESULT 6

CR584662 880 bp mRNA linear EST 26-JUL-2004  
 LOCUS CR584662 XGC-tailbud-head Xenopus tropicalis cDNA clone THDA042p12  
 DEFINITION 5', mRNA sequence.  
 ACCESSION CR584662.1 GI:50584662  
 VERSION CR584662  
 KEYWORDS EST.  
 SOURCE Xenopus tropicalis (western clawed frog)  
 ORGANISM Xenopus tropicalis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 Xenopodinae; Xenopus; Silurana.  
 1 (bases 1 to 880)  
 Croning,M.D.R., Ashurst,J.L., Taylor,R., Garrett,N. and Rogers,J.  
 Sanger Xenopus tropicalis EST project 2001 (2004)  
 TITLE Unpublished (2004)  
 JOURNAL  
 COMMENT Contact: Croning MDR  
 Sanger Institute  
 Hinxton, Cambridgeshire, CB10 1SA, UK  
 Email: trop@sanger.ac.uk  
 Sanger Xenopus tropicalis EST project 2001  
 TROPICALIS SEQUENCE ID: THDA042p12.plksp6  
 This sequence is from a Xenopus Gene Collection (XGC) library  
 constructed by Nigel Garrett.  
 Seq primer: SP6.  
 Location/Qualifiers  
 1..880  
 /organism="Xenopus tropicalis"

FEATURES  
source

Query Match 20.1%; Score 37.2; DB 5; Length 663;  
 Best Local Similarity 58.9%; Pred. No. 2.5;  
 Matches 63; Conservative 0; Mismatches 44; Indels 0; Gaps 0;  
 Qy 16 TGATCTATTTTCATTTGATCAAGGATTTATACAAAGAGAGACATCCATCCCGGTT 75  
 Db 345 TGTACTAGATAAAATATCCCAAGATTTTATAAGAACAAAGACTATTATGAGGAAT 404  
 Qy 76 AAGCAGTATGTTCCATCTAACAGAGAGGCTGCTGCACTGAGGAGG 122  
 Db 405 GAAGGAGTATGGCCCTGCTACACTTTGTGTCTGCACTGAGGAGG 451

Query Ma

RESULT_10	CNS017L2	1101 bp	DNA	linear	GSS 26-JUL-1999
LOCUS	Drosophila melanogaster	Genome survey sequence T7 end of BAC			
DEFINITION	BAC37D12 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL108176				
VERSION	AL108176.1	GI:5628480			
SOURCE	GSS.				
ORGANISM	Drosophila melanogaster (fruit fly)				

```

/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note=vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;

```



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 10, 2004, 15:57:32 ; Search time 391.831 Seconds  
(without alignments)  
9775.823 Million cell updates/sec

Title: US-09-909-207-27  
Perfect score: 81  
Sequence: 1 ATGAAATTGAAAGATTGAG.....CGGCTGTGCGGCTCATGCG 81

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_btg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sta.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	81	6	A42277 Sequence 27
2	81	100.0	81	6	A42283 Sequence 33
3	81	100.0	81	6	A48248 Sequence 27
4	81	100.0	81	6	A48249 Sequence 28
5	81	100.0	81	6	AR127042 Sequence
6	81	100.0	81	6	AR127047 Sequence
7	81	100.0	81	6	AR193072 Sequence
8	81	100.0	81	6	AR193073 Sequence
9	81	100.0	81	6	AR220046 Sequence
10	81	100.0	81	6	AR220051 Sequence
11	81	100.0	81	6	AR221329 Sequence
12	81	100.0	81	6	AR221334 Sequence
13	81	100.0	681	6	A42280 Sequence 30
14	81	100.0	681	6	A42282 Sequence 32
15	81	100.0	681	6	AR127045 Sequence
16	81	100.0	681	6	AR127046 Sequence
17	81	100.0	681	6	AR220049 Sequence
18	81	100.0	681	6	AR220050 Sequence
19	81	100.0	681	6	AR221332 Sequence

20	81	100.0	681	6	AR221333
21	81	100.0	1022	6	A42251
22	81	100.0	1022	6	A42285
23	81	100.0	1022	6	AR127019
24	81	100.0	1022	6	AR127049
25	81	100.0	1022	6	AR220023
26	81	100.0	1022	6	AR220053
27	81	100.0	1022	6	AR221306
28	81	100.0	1022	6	AR221336
29	79.4	98.0	684	1	AF220528
30	60.2	74.3	687	1	AY526092
31	60.2	74.3	1011	1	AF490981
32	60.2	74.3	1070	1	BPXINA
33	58.6	72.3	1789	1	AF326785
34	31.6	39.0	110000	8	CR380952_01
35	31.4	38.8	178061	2	AC108432
36	31.2	38.5	136956	2	AC137338
37	31.2	38.5	219065	2	AC137014
38	31.2	38.5	247840	2	AC130394
39	30.6	37.8	110000	2	BX572642_2
40	30.6	37.8	206200	9	AL160054
41	30.4	37.5	1735	5	AF528741
42	30.4	37.5	78483	9	AC098651
43	30.4	37.5	153414	9	AC093430
44	30.2	37.3	136175	8	AP004656
45	30.2	37.3	221534	2	AC111359

## ALIGNMENTS

RESULT 1  
A42277  
LOCUS A42277  
DEFINITION Sequence 27 from Patent EP0634490.  
ACCESSION A42277  
VERSION A42277.1 GI:2297765  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 81)  
AUTHORS De,B.E., Lahaya,A., Ledoux,P., Amory,A., Detroz,R., Andre,C. and Vetter,R.  
TITLE Xylanase derived from a bacillus species, expression vectors for such xylanase and other proteins, host organisms therefor and use thereof  
JOURNAL Patent: EP 0634490-A 27 18-JAN-1995;  
COMMENT SOLVAY (BE)  
Other publication NZ 260989 950828  
Other publication BR 9402834 950613  
Other publication JP 7067637 950314  
Other publication FI 943389 950116  
Other publication CA 2128050 950116  
Other publication AU 6743294 950127  
Other publication GB 2279955 950118.

FEATURES  
source  
1. 81

/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

## ORIGIN

Query Match 100.0%; Score 81; DB 6; Length 81;  
Best Local Similarity 100.0%; Pred. No. 2.4e-15;  
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATCAATTGAAAGATTGAGCGTGTGTTGTCATGTGTTGATGTTGTTGCTGCTGACACTG 60  
Db 1 ATCAATTGAAAGATTGAGCGTGTGTTGTCATGTGTTGATGTTGTTGCTGCTGACACTG 60  
Qy 61 ACGGCTGTGCGGCTCATGCG 81

[illegible]

DEFINITION	Sequence 27 from patent US 6180382.									
ACCESSION	AR127042									
VERSION	AR127042.1 GI:14113635									
KEYWORDS	Unknown.									
SOURCE	Unknown.									
ORGANISM	Unclassified.									
REFERENCE	1 (bases 1 to 81)									
AUTHORS	De Buyl,E., Lahaye,A., Ledoux,P., Amory,A., Detroz,R., Andre,C. and Vetter,R.									
TITLE	Xylanase derived from a bacillus species, expression vectors for such xylanase and other proteins, host organisms therefor and use thereof									
JOURNAL	Patent: US 6180382-A 27 30-JAN-2001;									
FEATURES	Location/Qualifiers									
source	1..81									
	/organism="unknown"									
	/mol_type="unassigned DNA"									
ORIGIN										
Query Match	100.0%; Score 81; DB 6; Length 81;									
Best Local Similarity	100.0%; Pred. No. 2.4e-15;									
Matches	81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1 ATGAATTTGAAAAGATTGAGGCTGTGTTTGTGATGTGATTTGGATTGTGCTGACACTG 60									
Db	1 ATGAATTTGAAAAGATTGAGGCTGTGTTTGTGATGTGATTTGGATTGTGCTGACACTG 60									
Qy	61 ACGGCTGTGCCGGCTCATGCG 81									
Db	61 ACGGCTGTGCCGGCTCATGCG 81									
RESULT 6										
AR127047										
LOCUS	AR127047 81 bp DNA linear PAT 16-MAY-2001									
DEFINITION	Sequence 33 from patent US 6180382.									
ACCESSION	AR127047									
VERSION	AR127047.1 GI:14113640									
KEYWORDS	Unknown.									
SOURCE	Unknown.									
ORGANISM	Unclassified.									
REFERENCE	1 (bases 1 to 81)									
AUTHORS	De Buyl,E., Lahaye,A., Ledoux,P., Amory,A., Detroz,R., Andre,C. and Vetter,R.									
TITLE	Xylanase derived from a bacillus species, expression vectors for such xylanase and other proteins, host organisms therefor and use thereof									
JOURNAL	Patent: US 6180382-A 33 30-JAN-2001;									
FEATURES	Location/Qualifiers									
source	1..81									
	/organism="unknown"									
	/mol_type="unassigned DNA"									
ORIGIN										
Query Match	100.0%; Score 81; DB 6; Length 81;									
Best Local Similarity	100.0%; Pred. No. 2.4e-15;									
Matches	81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1 ATGAATTTGAAAAGATTGAGGCTGTGTTTGTGATGTGATTTGGATTGTGCTGACACTG 60									
Db	1 ATGAATTTGAAAAGATTGAGGCTGTGTTTGTGATGTGATTTGGATTGTGCTGACACTG 60									
Qy	61 ACGGCTGTGCCGGCTCATGCG 81									
Db	61 ACGGCTGTGCCGGCTCATGCG 81									
RESULT 7										
AR193072										
LOCUS	AR193072 81 bp DNA linear PAT 20-APR-2002									
DEFINITION	Sequence 27 from patent US 6346407.									



AUTHORS	De,B.E., Lahaya,A., Ledoux,P., Amory,A., Detroz,R., Andre,C. and Vetter,R.
TITLE	Xylanase derived from a bacillus species, expression vectors for such xylanase and other proteins, host organisms therefor and use thereof
JOURNAL	Patent: EP 0634490-A 30 18-JAN-1995;
COMMENT	SOLVAY (BE) Other publication NZ 260989 950828 Other publication BR 9402834 950613 Other publication JP 7067637 950314 Other publication FI 943389 950116 Other publication CA 2128050 950116 Other publication NO 942652 950116 Other publication AU 6743294 950127 Other publication GB 2279955 950118.
FEATURES	Location/Qualifiers 1..681 /organism="unidentified" /mol_type="unassigned DNA" /db_xref="taxon:32644"
ORIGIN	
Query Match	100.0%; Score 81; DB 6; Length 681;
Best Local Similarity	100.0%; Pred. No. 2.3e-15;
Matches	81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGAATTTGAAAAGATTGAGGCTGTTTGTGATGTGATTGGATTGCTGACACTG 60 
Db	1 ATGAATTTGAAAAGATTGAGGCTGTTTGTGATGTGATTGGATTGCTGACACTG 60 
QY	61 ACGGCTGTGCCGGCTCATGCG 81 
Db	61 ACGGCTGTGCCGGCTCATGCG 81 
RESULT 15	
ARI27045	ARI27045 681 bp DNA linear PAT 16-MAY-2001
LOCUS	Sequence 30 from patent US 6180382.
DEFINITION	ARI27045
ACCESSION	ARI27045
VERSION	ARI27045.1 GI:14113638
KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 681)
AUTHORS	De Buy1,E., Lahaye,A., Ledoux,P., Amory,A., Detroz,R., Andre,C. and Vetter,R.
TITLE	Xylanase derived from a bacillus species, expression vectors for such xylanase and other proteins, host organisms therefor and use thereof
JOURNAL	Patent: US 6180382-A 30 30-JAN-2001;
FEATURES	Location/Qualifiers source 1..681 /organism="unknown" /mol_type="unassigned DNA"
ORIGIN	
Query Match	100.0%; Score 81; DB 6; Length 681;
Best Local Similarity	100.0%; Pred. No. 2.3e-15;
Matches	81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGAATTTGAAAAGATTGAGGCTGTTTGTGATGTGATTGGATTGCTGACACTG 60 
Db	1 ATGAATTTGAAAAGATTGAGGCTGTTTGTGATGTGATTGGATTGCTGACACTG 60 
QY	61 ACGGCTGTGCCGGCTCATGCG 81 
Db	61 ACGGCTGTGCCGGCTCATGCG 81 
Search completed:	November 10, 2004, 21:26:37
Job time :	394.831 secs

AUTHORS	De,B.E., Lahaya,A., Ledoux,P., Amory,A., Detroz,R., Andre,C. and Vetter,R.
TITLE	Xylanase derived from a bacillus species, expression vectors for such xylanase and other proteins, host organisms therefor and use thereof
JOURNAL	Patent: EP 0634490-A 32 18-JAN-1995;
COMMENT	SOLVAY (BE) Other publication NZ 260989 950828 Other publication BR 9402834 950613 Other publication JP 7067637 950314 Other publication FI 943389 950116 Other publication CA 2128050 950116 Other publication NO 942652 950116 Other publication AU 6743294 950127 Other publication GB 2279955 950118.
FEATURES	Location/Qualifiers 1..681 /organism="Bacillus pumilus" /mol_type="unassigned DNA" /strain="PRL B12" /db_xref="taxon:1408"
ORIGIN	
Query Match	100.0%; Score 81; DB 6; Length 681;
Best Local Similarity	100.0%; Pred. No. 2.3e-15;
Matches	81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 15:55:17 ; Search time 51.3814 Seconds  
(without alignments)  
8275.436 Million cell updates/sec

Title: US-09-909-207-27  
Perfect score: 81  
Sequence: 1 ATGAATTGAAAGATTGAG.....CGGCTGTCCGGCTCATGCG 81

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues 8269772  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 23Sep04: \*  
1: geneseqn1980s: \*  
2: geneseqn1990s: \*  
3: geneseqn2000s: \*  
4: geneseqn2001as: \*  
5: geneseqn2001bs: \*  
6: geneseqn2002as: \*  
7: geneseqn2002bs: \*  
8: geneseqn2003as: \*  
9: geneseqn2003bs: \*  
10: geneseqn2003cs: \*  
11: geneseqn2003ds: \*  
12: geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	81	2	AaQ73997 Xylanase
2	81	100.0	1022	2	AaQ80923 B. pumilu
3	79.4	98.0	684	12	Adj34969 DNA encod
4	30	37.0	1901	4	Ab117799 Drosophil
5	30	37.0	7142	4	Ab117798 Drosophil
6	29.4	36.3	508	10	Ad681488 Arabidops
7	29.4	36.3	1138	3	Aac52006 Arabidops
8	29.4	36.3	1216	3	Aac37584 Arabidops
9	28.6	35.3	1183	5	Abv29555 Human pro
10	28.6	35.3	1183	5	Abv23689 Human pro
11	28.4	35.1	1380	6	Abn66065 Streptoco
12	28.4	35.1	110000	6	Continuation (15 o
13	28.4	35.1	110000	6	Continuation (16 o
14	28	34.6	110000	10	Continuation (7 of
15	27.8	34.3	3623	4	Aa193584 Human pol
16	27.4	33.8	1905	12	Adp98598 C. albica
17	27.4	33.8	2790	1	Aan60473 Sequence
18	27.4	33.8	4106	10	Ad45341 Rat gene
19	27.4	33.8	4106	10	Ad56427 Rat gene
20	27.2	33.6	1332	10	Acf68431 Phototrab
21	27.2	33.6	102644	10	Acf65378 Phototrab

C	22	27.2	33.6	110000	10	ACF67367_11	Continuation (12 o
	23	27	33.3	4423	4	ABL27048	Ab127048 Drosophil
	24	26.6	32.8	345	6	ABN25404	Abn25404 Human ORF
	25	26.6	32.8	363	8	ACA35050	Aca35050 Prokaryot
	26	26.6	32.8	368	4	ABL14055	Ab114055 Drosophil
	27	26.6	32.8	5190	5	AAS69848	Aas69848 DNA encod
	28	26.6	32.8	5190	5	AAS81203	Aas81203 DNA encod
	29	26.6	32.8	5316	4	AAK51700	Aak51700 Human pol
	30	26.6	32.8	13033	4	ABL14054	Ab114054 Drosophil
	31	26.4	32.6	3234	12	ADP09637	Adp09637 Rice heat
	32	26.2	32.3	727	6	ABS77245	Ab577245 Frog heat
	33	26.2	32.3	1959	11	ADL72389	Adl72389 Cascor lo
	34	26.2	32.3	3915	4	ABL04067	Ab104067 Drosophil
	35	26.2	32.3	6147	3	AAZ53858	Aaz53858 Neisseria
	36	26.2	32.3	6197	4	ABL04066	Ab104066 Drosophil
	37	26.2	32.3	39746	4	ABL13398	Ab113398 Drosophil
	38	26.2	32.3	110000	3	AAF22303_1	Continuation (2 of
	39	26.2	32.3	110000	6	ABA90521_08	Continuation (9 of
	40	26	32.1	9741	6	ABL33323	Ade60106 Rat gene
	41	26	32.1	11049	6	ABL33323	Ab133323 Human imm
	42	26	32.1	11049	6	ABL32668	Ab132668 Human imm
	43	26	32.1	11049	6	ABL92218	Ab192218 Chemical
	44	26	32.1	11049	6	ABL49321	Ab149321 Human pol
	45	26	32.1	114693	8	AAD48308	Aad48308 Human tra

ALIGNMENTS

RESULT 1  
AAQ73997  
ID AAQ73997 standard; DNA; 81 BP.  
XX  
AC AAQ73997;  
XX  
DT 16-OCT-2003 (revised)  
DT 15-MAY-1996 (first entry)  
XX  
DE Xylanase gene presequence.  
XX  
KW Xylanase; thermostable enzyme; paper; pulp; bleaching; feedstuff; baking;  
KW promoter; signal peptide; ss.  
XX  
OS Bacillus pumilus; strain PRL B12 (ATCC 55443).  
XX  
PN AU9525086-A.  
XX  
PD 08-FEB-1996.  
XX  
PF 19-JUL-1995; 95AU-00025086.  
XX  
PR 26-JUL-1994; 94BE-00000706.  
PR 17-MAY-1995; 95BE-00000448.  
XX  
(SOLV ) SOLVAY SA.  
XX  
Pe Buyl E, Lahaye A, Ledoux P, Detroz R;  
XX  
WPI; 1996-117341/13.  
DR P-PSDB; AAR92055.  
XX  
Bacillus derived xylanase active over wide pH range - used in treatment  
of paper pulp, animal feeds and in bakery goods.  
XX  
Claim 27; Page 70; 94pp; English.  
XX  
The Bacillus pumilus PRL B12 xylanase gene presequence (AAQ73997)  
encoding the xylanase signal peptide (AAR92055) was incorporated along  
with the promoter sequence (AAQ73996) into vectors used for the  
expression of DNA (see AAT16101) coding for a mature, thermostable  
xylanase (AAR92053) of Bacillus sp. 720/1. The enzyme is produced in  
transformed hosts, pref. Bacillus licheniformis or B. pumilus, for use in  
the paper pulp, animal feed and baking industries. (Updated on 16-OCT-



```

CC 2003 to standardise OS field)
XX Sequence 81 BP; 16 A; 11 C; 26 G; 28 T; 0 U; 0 Other;
SQ

Query Match      100.0%; Score 81; DB 2; Length 81;
Best Local Similarity 100.0%; Pred. No. 2.5e-17;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAATTTGAAAAGATTGAGGCTGTTGTTGATGCTATTGATTTGCTGACACTG 60
Db 1 ATGAATTTGAAAAGATTGAGGCTGTTGTTGATGCTATTGATTTGCTGACACTG 60
Qy 61 ACGGCTGTGCGGCTCATGCG 81
Db 61 ACGGCTGTGCGGCTCATGCG 81

RESULT 2
AAQ80923
ID AAQ80923 standard; DNA; 1022 BP.
XX
AC AAQ80923;
DT 02-AUG-1995 (first entry)
XX
DE B. pumilus xylanase gene.
XX
KW Xylanase; endo-1,4-beta-D-xylanase; wood pulp; pulping; biobleaching;
KW bleaching; Bacillus licheniformis; ss.
XX
OS Bacillus pumilus.
XX
FH Key Location/Qualifiers
FT 5'UTR : 1..185
FT /tag= a
FT /note= "claim 12"
FT misc_difference 107
FT /tag= b
FT /note= "base n at position 107 is not identified in the
FT specification"
FT CDS 186..869
FT /tag= c
FT sig_peptide 186..266
FT /tag= d
FT /note= "claim 11"
FT mat_peptide 267..866
FT /tag= e
FT /EC number= "3.2.1.8"
FT /note= "claim 10"
XX
PN GB279955-A.
XX
XX 18-JAN-1995.
XX
XX 15-JUL-1993; 93GB-00014780.
XX
XX 15-JUL-1993; 93GB-00014780.
XX
XX (SOLV ) SOLVAY & CIE.
XX
XX Amory A, Andre C, De Buyl E, Detroz R, Lahaye A, Ledoux P;
XX
XX WPI; 1995-039214/06.
XX P-PSDB; AAR68849.
XX
XX Purified xylanase from Bacillus pumilus PRL B12 - esp. produced in
XX transformed Bacillus licheniformis, and related DNA, vectors, etc., used
XX for pre-treatment of wood pulp to reduce chlorine or ozone consumption in
XX subsequent bleaching.
XX
XX Claim 13; Fig 1a-1b; 97pp; English.
XX
XX A Bacillus pumilus PTL B12 (ATCC 55443) gene library was screened for

```

---

```

CC recombinant plasmids carrying the xylanase gene. A chromosomal fragment
CC obtained from isolate pBPX1 was subcloned and expressed in Escherichia
CC coli JM109. The sequence of the xylanase gene (as a 1022 bp Sau3AI
CC fragment) carried by a selected transformant is given in AAQ80923
XX
SQ Sequence 1022 BP; 329 A; 197 C; 229 G; 266 T; 0 U; 1 Other;
Query Match      100.0%; Score 81; DB 2; Length 1022;
Best Local Similarity 100.0%; Pred. No. 5.1e-17;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAATTTGAAAAGATTGAGGCTGTTGTTGATGCTATTGATTTGCTGACACTG 60
Db 186 ATGAATTTGAAAAGATTGAGGCTGTTGTTGATGCTATTGATTTGCTGACACTG 245
Qy 61 ACGGCTGTGCGGCTCATGCG 81
Db 246 ACGGCTGTGCGGCTCATGCG 266

RESULT 3
ADJ34969
ID ADJ34969 standard; DNA; 684 BP.
XX
AC ADJ34969;
XX
DT 22-APR-2004 (first entry)
XX
DE DNA encoding xylanase from an environmental sample seq id 185.
XX
KW antibacterial; fungicide; thermostable xylanase activity;
KW dough conditioning; beverage production; nutritional supplement;
KW animal feed; lignin reduction; wood product; xylan; bacterial infection;
KW fungal infection; coccidiosis; gene; ds.
XX
OS Unidentified.
XX
XX WO2003106654-A2.
XX
XX 24-DEC-2003.
XX
XX 16-JUN-2003; 2003WO-US019153.
XX
XX 14-JUN-2002; 2002US-0389299P.
XX
XX (DIVE-) DIVERSA CORP.
XX
XX Steer B, Callen W, Healey S, Hazlewood G, Wu D, Blum D;
XX Esteghlalian A;
XX
XX WPI; 2004-099016/10.
XX P-PSDB; ADJ34970.
XX
XX Novel xylanase recombinant polypeptide useful for improving textile
XX texture, treating paper, eliminating microorganisms.
XX
XX Claim 1; SEQ ID NO 185; 570pp; English.
XX
XX The invention describes an isolated or recombinant polypeptide (I),
XX having 50% or more identity to 190 300-1200 residue amino acid sequences
XX (SI), given in the specification, over a region of 100 or more residues
XX and the polypeptide as thermostable xylanase activity. (I) is useful for:
XX dough conditioning; beverage production; as a nutritional supplement in
XX animal feed; reducing lignin in a wood or a wood product; and for
XX eliminating and protecting animals from a microorganism comprising xylan.
XX The polynucleotide (II) encoding (I) is useful for amplifying nucleic
XX acid encoding a polypeptide having a xylanase activity which involves
XX amplification of a template nucleic acid with a primer pair capable of
XX amplifying (II) or its subsequence. (I) is useful for treating and
XX preventing bacterial infection and fungal infection e.g. coccidiosis.
XX This sequence encodes xylanase protein isolated from an environmental
XX sample.
XX

```

[illegible]

XX Arabidopsis thaliana expressed polynucleotide seq id 259.  
 DE genetically modified organism; transgenic organism; plant;  
 XX inhibitor testing; activator testing; modifier testing; fungicide;  
 KW insecticide; genetic function; genetic regulation; cellular metabolism;  
 KW gene; ss.  
 XX Arabidopsis thaliana.  
 OS US2003115639-A1.  
 XX 19-JUN-2003.  
 XX 26-JAN-2001; 2001US-00770961.  
 XX 27-JAN-2000; 2000US-0178466P.  
 XX (GORL/) GORLACH J.  
 PA (ANY/) AN Y.  
 PA (HAMI/) HAMILTON C M.  
 PA (PRIC/) PRICE J L.  
 PA (RAIN/) RAINES T M.  
 PA (YUY/) YU Y.  
 PA (RAME/) RAMEKA J G.  
 PA (PAGE/) PAGE A.  
 PA (MATH/) MATHAW A V.  
 PA (LEDF/) LEDFORD B L.  
 PA (WOES/) WOESSNER J P.  
 PA (HAAS/) HAAS W D.  
 PA (GARC/) GARCIA C A.  
 PA (KRIC/) KRICKER M.  
 PA (SLAT/) SLATER T.  
 PA (DAVI/) DAVIS K R.  
 PA (ALLE/) ALLEN K.  
 PA (HOFF/) HOFFMAN N.  
 PA (HURB/) HURBAN P.  
 XX Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y,  
 PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;  
 PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;  
 PI Hurban P;  
 XX WPI; 2003-810930/76.  
 XX Novel Arabidopsis thaliana nucleic acids useful for generating  
 PT genetically modified transgenic organisms, for screening biologically  
 PT active agents such as fungicides, insecticides.  
 XX Claim 1; SEQ ID NO 259; 44pp; English.  
 CC The invention describes a nucleic acid (I) comprising a sequence capable  
 CC of hybridising under stringent conditions to any one of 999 fully defined  
 CC Arabidopsis thaliana sequences (I) as given in specification e.g., 360,  
 CC 1137, 455, 219, 472 nucleotides, etc, or its fragment. (I) is useful as a  
 CC hybridisation probe to complementary molecules in a cDNA library. (I) is  
 CC also useful for generating genetically modified and transgenic organisms,  
 CC usually plant cells and plants. A protein encoded by (I) is useful in  
 CC screening assays to determine the effect of candidate inhibitors,  
 CC activators or modifiers of the gene product. The protein is also useful  
 CC for screening biologically active agents e.g., fungicides and  
 CC insecticides. A genetically modified cell, comprising an exogenous  
 CC nucleic acid, where the nucleic acid comprises transcription regulatory  
 CC sequences operably linked to a sequence capable of hybridising under  
 CC stringent conditions to (I) is useful in the study of genetic function  
 CC and regulation, for alteration of the cellular metabolism and for  
 CC screening compounds that may affect the biological function of the gene  
 CC or gene product. This sequence represents an Arabidopsis thaliana  
 CC polynucleotide of the invention.  
 XX SQ Sequence 508 BP; 165 A; 109 C; 90 G; 144 T; 0 U; 0 Other;  
 Query Match 36.3%; Score 29.4; DB 10; Length 508;

Best Local Similarity 70.9%; Pred. No. 6.9;  
 Matches 39; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
 Qy 1 ATGATTTGAAAGATTGAGCTGCTTTTGTGATGTGATTTGCTGA 55  
 Db 159 ATGCATGTGAAAGCTTGGATGATGCTTTTGTGATTTGCTGA 105  
 RESULT 7  
 AAC52006  
 ID AAC52006 standard; DNA; 1138 BP.  
 XX AC AAC52006;  
 XX 18-OCT-2000 (first entry)  
 XX Arabidopsis thaliana DNA fragment SEQ ID NO: 70190.  
 XX Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway; metabolic pathway;  
 KW promoter; termination sequence; ss.  
 XX Arabidopsis thaliana.  
 PN EP1033405-A2.  
 PD 06-SEP-2000.  
 XX 25-FEB-2000; 2000EP-00301439.  
 XX 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 28-APR-1999; 99US-0130891P.  
 PR 30-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 04-MAY-1999; 99US-0132407P.  
 PR 05-MAY-1999; 99US-0132484P.  
 PR 06-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 07-MAY-1999; 99US-0132487P.  
 PR 11-MAY-1999; 99US-0132863P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.

[illegible]

QY  
1 ATGAATTGAAAAGATTGAGCGCTGTTGTGGTATGTGTAATGGATTGTGCTGA 55

n6  
956 ATGCCAAGTCGTAAGAAGCTTGACATGATGCTTTGTTGGTCTTGGATTGTGCTGA 1010

RESULT 8  
AAC37584  
ID AAC37584 standard; DNA; 1216 BP.  
XX  
AC AAC37584;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 17922.  
XX  
KW Hybridisation assay; Genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0128845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.

PR	20-AUG-1999;	99US-0143723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151920P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0156559P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	14-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Query Match

Best Local Similarity

Matches 39; Conservative

Score 29.4; DB 3; Length 1216;

Pred. No. 8.8;

0; Mismatches 16; Indels 0; Gaps

Qy	1	ATCAATTGAAAGATGAGCGCTTGTCTTGTGATGATGATTCGATTGCTGTA	55
Db	1012	ATCGATGTGTAAAGCTTGATGATGATGCTTTTGTGTGCTTCGATTGCTGTA	1066

RESULT 9

ABV29555/c

ID ABV29555 standard; cDNA; 1183 BP.

XX

AC

XX

ABV29555;

XX





	Query Match	35.1%	Score 28.4;	DB 6;	Length 110000;	
	Best Local Similarity	70.4%;	Pred. No. 67;			
	Matches 38; Conservative	0;	Mismatches 16;	Indels 0;	Gaps 0;	
QY	1	ATCAATTGAAACAGATTGAGCGCTCTTTGTGTGATGTGATTGGATTGTGCTG	54			
Db	103038	ATCAACGTTATTCCCATTTGAGTCITTTTTACAGTGATGCTTATTGGATTGTGTCG	102985			

b		73928 GCGGCAC 73935
	RESULT 15	
	AAI93584	
D	AAI93584 standard; cDNA; 3623 BP.	
X	C	
X	AAI93584;	
X	T	
X	06-NOV-2001 (first entry)	
E	Human polynucleotide SEQ ID NO 13644 .	
X	Human; cytokine; cell proliferation; cell differentiation; gene therapy;	
X	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;	
W	tissue growth factor; immunomodulatory; cancer; leukaemia;	
W	nervous system disorders; arthritis; inflammation; ss.	
X		
X	Homo sapiens.	
S	::	
X	WO200164835-A2.	
N		
N		
X		
X	07-SEP-2001.	
X		
X	26-FEB-2001; 2001WO-US004927.	
F		
X		
X	28-FEB-2000; 2000US-00515126.	
R		

X  
X  
X  
X  
X  
R

(HYSE-) HYSEQ INC.  
  
Tang YT, Liu C, Drmanac RT;  
WFI; 2001-514838/56.  
P-PSDB; AAO13653.

and treating e.g. leukemia, inflammation and immune disorders.

Claim 1; SEQ ID NO 13644; 1399pp + Sequence Listing; English.

The invention relates to human polynucleotides (AA179941-AA193841) and the encoded proteins (AA000010-AA013910) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)

Sequence 3623 BP; 1143 A; 501 C; 572 G; 1303 T; 0 U; 104 Other;

Query Match 34.3%; Score 27.8; DB 4; Length 3623;  
Best Local Similarity 65.1%; Pred. No. 41;  
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

6 TTTGAAAGATTGAGCGTGTGTTGTCATGTGTTATTCGATTGTGCTGCACACTGACGGC 65

Db 2724 TTGATATAGATATATGATATTGATTCTCGAGCTGCTTTCTGTAATTTTTCTGACTTTGGCATC 2783

Qy 66 TGT 68

Db 2784 TGT 2786

Search completed: November 10, 2004, 17:09:40  
Job time : 54.3814 secs

GenCore version 5.1.6  
Copyright (C) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 17:09:58 ; Search time 9.30509 Seconds  
(without alignments)  
6187.359 Million cell updates/sec

Title: US-09-909-207-27  
Perfect score: 81  
Sequence: 1 ATGAATTTGAAAGATTGAG.....CGCTGTGCGGCTCATGG 81

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	81	3	US-08-275-526C-27
2	81	100.0	81	3	US-08-275-526C-33
3	81	100.0	81	3	US-08-470-953A-27
4	81	100.0	81	3	US-08-470-953A-28
5	81	100.0	81	4	US-09-076-677-27
6	81	100.0	81	4	US-09-076-677-33
7	81	100.0	81	4	US-09-073-055-27
8	81	100.0	81	4	US-09-073-055-33
9	81	100.0	681	3	US-08-275-526C-30
10	81	100.0	681	3	US-08-275-526C-32
11	81	100.0	681	4	US-09-076-677-30
12	81	100.0	681	4	US-09-076-677-32
13	81	100.0	681	4	US-09-073-055-30
14	81	100.0	681	4	US-09-073-055-32
15	81	100.0	1022	3	US-08-275-526C-1
16	81	100.0	1022	3	US-08-275-526C-35
17	81	100.0	1022	4	US-09-076-677-1
18	81	100.0	1022	4	US-09-076-677-35
19	81	100.0	1022	4	US-09-073-055-1
20	81	100.0	1022	4	US-09-073-055-35
21	28.4	35.1	987	4	US-09-270-767-13499
22	27.8	34.3	494	4	US-09-621-976-2247
23	27.4	33.8	408	4	US-09-248-796A-2584
24	26.2	32.3	324	4	US-09-614-221A-198
25	26.2	32.3	726	4	US-09-248-796A-9223
26	26	32.1	11049	4	US-10-204-708-21
27	25.4	31.4	262	4	US-09-313-294A-1889

c	28	25.4	31.4	1317	4	US-09-248-796A-5112	Sequence 5112, Ap
c	29	25.2	31.1	1710	3	US-09-068-960-14	Sequence 14, Appl
c	30	25	30.9	564	4	US-09-248-796A-596	Sequence 596, Appl
c	31	25	30.9	1377	4	US-09-134-000C-3078	Sequence 3078, Ap
c	32	25	30.9	2013	4	US-09-543-681A-1807	Sequence 1807, Ap
c	33	25	30.9	4487	4	US-09-484-970B-96	Sequence 96, Appl
c	34	25	30.9	36412	4	US-08-311-731A-132	Sequence 132, Appl
c	35	24.8	30.6	849	4	US-09-248-796A-7183	Sequence 7183, Ap
c	36	24.8	30.6	1230	4	US-09-248-796A-1393	Sequence 1393, Ap
c	37	24.8	30.6	1963	5	PCT-US91-07715A-1	Sequence 1, Appl
c	38	24.8	30.6	1970	1	US-07-596-867C-1	Sequence 1, Appl
c	39	24.8	30.6	1664976	4	US-08-916-421B-1	Sequence 1, Appl
c	40	24.8	30.6	1664976	4	US-09-692-570-1	Sequence 1, Appl
c	41	24.6	30.4	1513	4	US-09-301-666A-9	Sequence 9, Appl
c	42	24.6	30.4	1513	4	US-09-301-217-9	Sequence 9, Appl
c	43	24.6	30.4	2106	4	US-09-248-796A-2525	Sequence 2525, Ap
c	44	24.4	30.1	1236	4	US-09-248-796A-6578	Sequence 6578, Ap
c	45	24.4	30.1	2542	3	US-08-941-445A-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1  
US-08-275-526C-27  
; Sequence 27, Application US/08275526C  
; Patent No. 6180382  
; GENERAL INFORMATION:  
; APPLICANT: DE BUYL, ERIC  
; APPLICANT: LAHAYE, ANDR E  
; APPLICANT: LEDOUX, PIERRE  
; APPLICANT: AMORY, ANTOINE  
; APPLICANT: DETROZ, REN  
; APPLICANT: ANDRE, CHRISTOPHE  
; APPLICANT: VETTER, ROMAN  
; TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,  
; TITLE OF INVENTION: EXPRESSION VECTORS FOR SUCH XYLANASE AND  
; TITLE OF INVENTION: OTHER PROTEINS, HOST ORGANISMS THEREFOR AND  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.  
; STREET: 2000 K St., N.W., Suite 200  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/275,526C  
; FILING DATE: 15-JUL-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gadiano, Wilhem F.  
; REGISTRATION NUMBER: 37,136  
; REFERENCE/DOCKET NUMBER: 4121-49  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 429-0625  
; TELEFAX: (202) 293-0625  
; TELEX: 650 383 5605  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 81 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: genomic DNA  
US-08-275-526C-27





```

; STREET: 2000 K St., N.W., Suite 200
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/076,677
; FILING DATE: 12-May-1998
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/275,526
; FILING DATE: 15-JUL-1994
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Gadiano, Wilhem F.
; REGISTRATION NUMBER: 37,136
; REFERENCE/DOCKET NUMBER: 4121-49
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 429-0625
; TELEFAX: (202) 293-0625
;
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 81 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Bacillus pumilus
; STRAIN: PRL B12
; SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-09-076-677-33

Query Match 100.0%; Score 81; DB 4; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.8e-19;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAATTGAAAGATTGAGGCTGTTGTTGTGATGTGATTTGGATTGTGCTGACACTG 60
Db 1 ATGAATTGAAAGATTGAGGCTGTTGTTGTGATGTGATTTGGATTGTGCTGACACTG 60

QY 61 ACGGCTGCCGGCTCATCG 81
Db 61 ACGGCTGCCGGCTCATCG 81

RESULT 7
US-09-073-055-27
; Sequence 27, Application US/09073055
; Patent No. 6426211
;
; GENERAL INFORMATION:
; APPLICANT: DE BUYL, ERIC
; LAHAYE, ANDR E
; LEDOUX, PIERRE
; AMORY, ANTOINE
; DETROZ, REN
; ANDRE, CHRISTOPHE
; VETTER, ROMAN
;
; TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,
; EXPRESSION VECTORS FOR SUCH XYLANASE AND
; OTHER PROTEINS, HOST ORGANISMS THEREFOR AND
; USE THEREOF
;
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.
; STREET: 2000 K St., N.W., Suite 200
; CITY: Washington
; STATE: D.C.
;
; COUNTRY: U.S.A.
; ZIP: 20006
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,055
; FILING DATE: 05-May-1998
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/275,526
; FILING DATE: 15-JUL-1994
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Gadiano, Wilhem F.
; REGISTRATION NUMBER: 37,136
; REFERENCE/DOCKET NUMBER: 4121-49
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 429-0625
; TELEFAX: (202) 293-0625
;
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 81 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Bacillus pumilus
; STRAIN: PRL B12
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-073-055-27

Query Match 100.0%; Score 81; DB 4; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.8e-19;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAATTGAAAGATTGAGGCTGTTGTTGTGATGTGATTTGGATTGTGCTGACACTG 60
Db 1 ATGAATTGAAAGATTGAGGCTGTTGTTGTGATGTGATTTGGATTGTGCTGACACTG 60

QY 61 ACGGCTGCCGGCTCATCG 81
Db 61 ACGGCTGCCGGCTCATCG 81

RESULT 8
US-09-073-055-33
; Sequence 33, Application US/09073055
; Patent No. 6426211
;
; GENERAL INFORMATION:
; APPLICANT: DE BUYL, ERIC
; LAHAYE, ANDR E
; LEDOUX, PIERRE
; AMORY, ANTOINE
; DETROZ, REN
; ANDRE, CHRISTOPHE
; VETTER, ROMAN
;
; TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,
; EXPRESSION VECTORS FOR SUCH XYLANASE AND
; OTHER PROTEINS, HOST ORGANISMS THEREFOR AND
; USE THEREOF
;
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.
; STREET: 2000 K St., N.W., Suite 200
; CITY: Washington
; STATE: D.C.
;
; COUNTRY: U.S.A.
; ZIP: 20006
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/076,677
; FILING DATE: 12-May-1998
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/275,526
; FILING DATE: 15-JUL-1994
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Gadiano, Wilhem F.
; REGISTRATION NUMBER: 37,136
; REFERENCE/DOCKET NUMBER: 4121-49
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 429-0625
; TELEFAX: (202) 293-0625
;
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 81 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Bacillus pumilus
; STRAIN: PRL B12
; SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-09-076-677-33

Query Match 100.0%; Score 81; DB 4; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.8e-19;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAATTGAAAGATTGAGGCTGTTGTTGTGATGTGATTTGGATTGTGCTGACACTG 60
Db 1 ATGAATTGAAAGATTGAGGCTGTTGTTGTGATGTGATTTGGATTGTGCTGACACTG 60

QY 61 ACGGCTGCCGGCTCATCG 81
Db 61 ACGGCTGCCGGCTCATCG 81

RESULT 7
US-09-073-055-27
; Sequence 27, Application US/09073055
; Patent No. 6426211
;
; GENERAL INFORMATION:
; APPLICANT: DE BUYL, ERIC
; LAHAYE, ANDR E
; LEDOUX, PIERRE
; AMORY, ANTOINE
; DETROZ, REN
; ANDRE, CHRISTOPHE
; VETTER, ROMAN
;
; TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,
; EXPRESSION VECTORS FOR SUCH XYLANASE AND
; OTHER PROTEINS, HOST ORGANISMS THEREFOR AND
; USE THEREOF
;
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.
; STREET: 2000 K St., N.W., Suite 200
; CITY: Washington
; STATE: D.C.
;
; COUNTRY: U.S.A.
; ZIP: 20006
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,055
; FILING DATE: 05-May-1998
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/275,526
; FILING DATE: 15-JUL-1994
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Gadiano, Wilhem F.
; REGISTRATION NUMBER: 37,136
; REFERENCE/DOCKET NUMBER: 4121-49
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 429-0625
; TELEFAX: (202) 293-0625
;
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 81 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Bacillus pumilus
; STRAIN: PRL B12
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-073-055-27
```



```
;
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 681 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORGANISM: Bacillus pumilus
; STRAIN: PRL B12
; US-08-275-526C-32

Query Match          100.0%; Score 81; DB 3; Length 681;
Best Local Similarity 100.0%; Pred. No. 3.8e-19;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAATTTGAAAAGATTGAGGCTGTTGTTGTGATGTGATTTGGATTGTGCTGACACTG 60
Db 1 ATGAATTTGAAAAGATTGAGGCTGTTGTTGTGATGTGATTTGGATTGTGCTGACACTG 60

QY 61 ACGGCTGTGCCGCTCATCG 81
Db 61 ACGGCTGTGCCGCTCATCG 81

RESULT 11
US-09-076-677-30
; Sequence 30, Application US/09076677
; Patent No. 6423523
; GENERAL INFORMATION:
; APPLICANT: DE BUYL, ERIC
; LAHAYE, ANDREE
; LEDOUX, PIERRE
; AMORY, ANTOINE
; DETROZ, RENE
; ANDRE, CHRISTOPHE
; VETTER, ROMAN
; TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,
; EXPRESSION VECTORS FOR SUCH XYLANASE AND
; OTHER PROTEINS, HOST ORGANISMS THEREFOR AND
; USE THEREOF
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WILLIAN BRINKS HOFER GILSON & LIONE, P.C.
; STREET: 2000 K St., N.W., Suite 200
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/076,677
; FILING DATE: 12-May-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/275,526
; FILING DATE: 15-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Gadiano, Wilhem F.
; REGISTRATION NUMBER: 37,136
; REFERENCE/DOCKET NUMBER: 4121-49
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 429-0625
; TELEFAX: (202) 293-0625
; TELEX: 650 383 5605
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 681 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORGANISM: Bacillus pumilus
; STRAIN: PRL B12
; US-09-076-677-30

;
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 30:

Query Match          100.0%; Score 81; DB 4; Length 681;
Best Local Similarity 100.0%; Pred. No. 3.8e-19;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAATTTGAAAAGATTGAGGCTGTTGTTGTGATGTGATTTGGATTGTGCTGACACTG 60
Db 1 ATGAATTTGAAAAGATTGAGGCTGTTGTTGTGATGTGATTTGGATTGTGCTGACACTG 60

QY 61 ACGGCTGTGCCGCTCATCG 81
Db 61 ACGGCTGTGCCGCTCATCG 81

RESULT 12
US-09-076-677-32
; Sequence 32, Application US/09076677
; Patent No. 6423523
; GENERAL INFORMATION:
; APPLICANT: DE BUYL, ERIC
; LAHAYE, ANDREE
; LEDOUX, PIERRE
; AMORY, ANTOINE
; DETROZ, RENE
; ANDRE, CHRISTOPHE
; VETTER, ROMAN
; TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,
; EXPRESSION VECTORS FOR SUCH XYLANASE AND
; OTHER PROTEINS, HOST ORGANISMS THEREFOR AND
; USE THEREOF
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WILLIAN BRINKS HOFER GILSON & LIONE, P.C.
; STREET: 2000 K St., N.W., Suite 200
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/076,677
; FILING DATE: 12-May-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/275,526
; FILING DATE: 15-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Gadiano, Wilhem F.
; REGISTRATION NUMBER: 37,136
; REFERENCE/DOCKET NUMBER: 4121-49
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 429-0625
; TELEFAX: (202) 293-0625
; TELEX: 650 383 5605
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 681 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORGANISM: Bacillus pumilus
; STRAIN: PRL B12
```



Qy 1 ATGAATTTGAAAAGATTGAGGCTGTTTGTGATGATGTAATTGGATTTGTGCTGACACTG 60

Db 1 ATGAATTTGAAAAGATTGAGGCTGTTGTTGTGATGATGATTGGATTGTTGCTGACACTG 60  
QY 61 ACGGCTGTGCCGCTCATGCG 81  
Db 61 ACGGCTGTGCCGCTCATGCG 81

## RESULT 15

US-08-275-526C-1  
; Sequence 1, Application US/08275526C  
; Patent No. 6180382  
; GENERAL INFORMATION:  
; APPLICANT: DE BUYL, ERIC  
; APPLICANT: LAHAYE, ANDR E  
; APPLICANT: LEDOUX, PIERRE  
; APPLICANT: AMORY, ANTOINE  
; APPLICANT: DETROZ, REN  
; APPLICANT: ANDRE, CHRISTOPHE  
; APPLICANT: VETTER, ROMAN  
; TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,  
; TITLE OF INVENTION: EXPRESSION VECTORS FOR SUCH XYLANASE AND,  
; TITLE OF INVENTION: OTHER PROTEINS, HOST ORGANISMS THEREFOR AND  
; TITLE OF INVENTION: USE THEREOF  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WILLIAN BRINKS HOFER GILSON & LIONE, P.C.  
; STREET: 2000 K St., N.W., Suite 200  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/275,526C  
; FILING DATE: 15-JUL-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gadiano, Wilhem F.  
; REGISTRATION NUMBER: 37,136  
; REFERENCE/DOCKET NUMBER: 4121-49  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 429-0625  
; TELEFAX: (202) 293-0625  
; TELEX: 650 383 5605  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1022 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: genomic DNA  
; ORIGINAL SOURCE:  
; ORGANISM: Bacillus pumilus  
; STRAIN: PRL B12  
; US-08-275-526C-1

Query Match 100.0%; Score 81; DB 3; Length 1022;  
Best Local Similarity 100.0%; Pred. No. 4.3e-19;  
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAATTTGAAAAGATTGAGGCTGTTGTTGTGATGATGATTGGATTGTTGCTGACACTG 60  
Db 186 ATGAATTTGAAAAGATTGAGGCTGTTGTTGTGATGATGATTGGATTGTTGCTGACACTG 245  
QY 61 ACGGCTGTGCCGCTCATGCG 81  
Db 246 ACGGCTGTGCCGCTCATGCG 266

Search completed: November 11, 2004, 01:34:29  
Job time : 11.3051 secs

GenCore version 5.1.6  
Copyright (C) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 21:26:43 ; Search time 51.7881 Seconds  
(without alignments)  
8421.106 Million cell updates/sec

Title: US-09-909-207-27

Perfect score: 81

Sequence: 1 ATGATTTTCAAGATTGAG.....CGGCTGTGCGGCTCATGCG 81

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3611042 seqs, 2692057975 residues

Total number of hits satisfying chosen parameters: 7222084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/prodata/1/pubpna/PCT05\_PUBCOMB.seq:\*
- 7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/prodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/prodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/prodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	81	9	US-09-909-207-27
2	81	100.0	81	9	US-09-909-207-28
3	29.4	36.3	508	10	Sequence 259, App
4	29	35.8	2363	13	US-10-027-632-263227
5	29	35.8	2363	13	US-10-027-632-263228
6	29	35.8	2363	15	US-10-027-632-263227
7	29	35.8	2363	15	US-10-027-632-263228
8	28.8	35.6	233060	13	US-10-087-192-97
9	28	34.6	72314	13	US-10-087-192-1615
10	28	34.6	87731	13	US-10-087-192-1342
11	27.6	34.1	614	18	US-10-425-115-162687
12	27.6	34.1	1383	18	US-10-425-115-137066

C 13	27.4	33.8	263	18	US-10-425-115-83971
C 14	27.4	33.8	648	17	US-10-437-963-16223
C 15	27.4	33.8	1763	17	US-10-437-963-71944
C 16	27.2	33.6	1128	18	US-10-425-115-60737
C 17	27.2	33.6	193853	13	US-10-087-192-1663
C 18	27	33.3	681	15	US-10-369-493-41153
C 19	26.8	33.1	343	16	US-10-424-599-24861
C 20	26.8	33.1	454	17	US-10-767-701-9260
C 21	26.8	33.1	5755	18	US-10-739-930-210
C 22	26.6	32.8	363	16	US-10-282-122A-22920
C 23	26.4	32.6	2274	16	US-10-424-599-109383
C 24	26.2	32.3	324	18	US-10-793-639-198
C 25	26.2	32.3	727	9	US-09-910-943-478
C 26	26.2	32.3	990	16	US-10-425-114-3098
C 27	26.2	32.3	1052	18	US-10-425-115-66655
C 28	26.2	32.3	1278	16	US-10-425-114-35562
C 29	26.2	32.3	1681	18	US-10-425-115-76243
C 30	26.2	32.3	1959	16	US-10-410-031-161
C 31	26	32.1	401	13	US-10-027-632-163378
C 32	26	32.1	401	15	US-10-027-632-163378
C 33	26	32.1	404	16	US-10-424-599-1622
C 34	26	32.1	1011	16	US-10-425-114-9394
C 35	26	32.1	1932	16	US-10-425-114-11077
C 36	26	32.1	2160	18	US-10-425-115-59754
C 37	26	32.1	2354	16	US-10-424-599-129083
C 38	26	32.1	9741	15	US-10-311-455-1296
C 39	26	32.1	11049	15	US-10-204-708-21
C 40	26	32.1	11049	15	US-10-311-455-641
C 41	26	32.1	11049	16	US-10-240-589C-27
C 42	25.8	31.9	2896	15	US-10-128-714-352
C 43	25.8	31.9	2897	15	US-10-128-714-352
C 44	25.8	31.9	248436	13	US-10-087-192-2014
C 45	25.8	31.9	367378	15	US-10-312-841-1

ALIGNMENTS

RESULT 1

US-09-909-207-27

Sequence 27, Application US/09909207

Patent No. US20020115181A1

GENERAL INFORMATION:

APPLICANT: ANDREE LAHAYE

ERIC DE BUYL

PIERRE LEDOUX

RENE DETROZ

TITLE OF INVENTION: Xylanase, microorganisms produced it,  
DNA molecule, processes for preparation of this xylanase  
and uses thereof

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE

STREET: 2000 K St., N.W., Suite 200

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION NUMBER: US/09/909,207

APPLICATION NUMBER: 19-Jul-2001

FILING DATE: 19-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/470,953

FILING DATE: 06-JUNE-1995

ATTORNEY/AGENT INFORMATION:

NAME: Willem F. Gadiano, Esq.

REGISTRATION NUMBER: 37,136

```

; REFERENCE/DOCKET NUMBER: 4121-40
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-429-0625
; TELEFAX: (202) 293-1850
; TELEX: 650 383-5605
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 81 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-909-207-27

Query Match      100.0%; Score 81; DB 9; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.1e-17;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAATTGAAAAGATTGAGGCTGTGTTGTGATGCTGATTGCGTCTGCTGACACTG 60
Db 1 ATGAATTGAAAAGATTGAGGCTGTGTTGTGATGCTGATTGCGTCTGCTGACACTG 60

QY 61 ACGGCTGTCCGGCTCATCG 81
Db 61 ACGGCTGTCCGGCTCATCG 81

RESULT 2
US-09-909-207-28
; Sequence 28, Application US/09909207
; Patent No. US20020115181A1
; GENERAL INFORMATION:
; APPLICANT: ANDREE LAHAYE
; ERIC DE BUYL
; PIERRE LEDOUX
; RENE DETROZ
; TITLE OF INVENTION: Xylanase, microorganisms produced it,
; DNA molecule, processes for preparation of this xylanase
; and uses thereof
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSER: WILLIAM BRINKS HOPFER GILSON & LIONE
; STREET: 2000 K St., N.W., Suite 200
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/909,207
; FILING DATE: 19-Jul-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/470,953
; FILING DATE: 06-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilhem F. Gadiano, Esq.
; REGISTRATION NUMBER: 37,136
; REFERENCE/DOCKET NUMBER: 4121-40
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-429-0625
; TELEFAX: (202) 293-1850
; TELEX: 650 383-5605
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 81 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..81
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-909-207-28

Query Match      100.0%; Score 81; DB 9; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.1e-17;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAATTGAAAAGATTGAGGCTGTGTTGTGATGCTGATTGCGTCTGCTGACACTG 60
Db 1 ATGAATTGAAAAGATTGAGGCTGTGTTGTGATGCTGATTGCGTCTGCTGACACTG 60

QY 61 ACGGCTGTCCGGCTCATCG 81
Db 61 ACGGCTGTCCGGCTCATCG 81

RESULT 3
US-09-770-961-259/c
; Sequence 259, Application US/09770961
; Publication No. US20030115639A1
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Kameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hubban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2026 (PARA-015PRV)
; CURRENT APPLICATION NUMBER: US/09/770,961
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,466
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 259
; LENGTH: 508
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-961-259

Query Match      36.3%; Score 29.4; DB 10; Length 508;
Best Local Similarity 70.9%; Pred. No. 6.5;
Matches 39; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGAATTGAAAAGATTGAGGCTGTGTTGTGATGCTGATTGCGTCTGCTGACACTG 55
Db 159 ATGCATGTGTAAAGCTTGCATGATGCTTTGTTGTGCTCTGGATTGCTGCTG 105

RESULT 4
US-10-027-632-263227
; Sequence 263227, Application US/10027632
; Publication No. US20020198371A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263227
; LENGTH: 2363
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-263227

Query Match 35.8%; Score 29; DB 13; Length 2363;
Best Local Similarity 67.2%; Pred. No. 15; Gaps 0;
Matches 41; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGAATTCGAAAGATTGAGCGCTGTTTGTGATGTATGGAATTGTGCTGACACTG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1219 ATGCATCTTATTAGATGAGAAGTTGTAGATATGATATATGTTTGTCTGAGATTG 1278

QY 61 A 61
    |
Db 1279 A 1279

RESULT 5
US-10-027-632-263228
; Sequence 263228, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263228
; LENGTH: 2363
; TYPE: DNA
; ORGANISM: Human

```

;; PRIOR FILING DATE: 2000-07-12  
;; PRIOR APPLICATION NUMBER: US 60/198,676  
;; PRIOR FILING DATE: 2000-04-20  
;; PRIOR APPLICATION NUMBER: US 60/193,483  
;; PRIOR FILING DATE: 2000-03-29  
;; PRIOR APPLICATION NUMBER: US 60/185,218  
;; PRIOR FILING DATE: 2000-02-24  
;; PRIOR APPLICATION NUMBER: US 60/167,363  
;; PRIOR FILING DATE: 1999-11-23  
;; PRIOR APPLICATION NUMBER: US 60/156,358  
;; PRIOR FILING DATE: 1999-09-28  
;; PRIOR APPLICATION NUMBER: US 60/146,002  
;; PRIOR FILING DATE: 1999-08-09  
;; NUMBER OF SEQ ID NOS: 325720  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 263228  
;; LENGTH: 2363  
;; TYPE: DNA  
;; ORGANISM: Human  
US-10-027-632-263228

Query Match 35.8%; Score 29; DB 15; Length 2363;  
Best Local Similarity 67.2%; Pred. No. 15;  
Matches 41; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
QY 1 ATGAATTTCAAGAGATTGAGGCTGTTGTTGGATGTTGATGTTGATGTTGCTGACACTG 60  
DB 1219 ATGCATCTTATTAGAATGAGATGTTGTAGATAATGATTATATGTTGTTGCTCTGAGATTG 1278  
QY 61 A 61  
DB 1279 A 1279

RESULT 8  
US-10-087-192-97  
;; Sequence 97, Application US/10087192  
;; Publication No. US20020182586A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Morris, David W.  
;; APPLICANT: Engelhard, Eric K.  
;; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
;; TITLE OF INVENTION: CANCER  
;; FILE REFERENCE: 529452000122  
;; CURRENT APPLICATION NUMBER: US/10/087,192  
;; CURRENT FILING DATE: 2002-03-01  
;; PRIOR APPLICATION NUMBER: US 09/747,377  
;; PRIOR FILING DATE: 2000-12-22  
;; PRIOR APPLICATION NUMBER: US 09/798,586  
;; PRIOR FILING DATE: 2001-03-02  
;; NUMBER OF SEQ ID NOS: 2059  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 97  
;; LENGTH: 233060  
;; TYPE: DNA  
;; ORGANISM: Mus musculus  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: (1)...(233060)  
;; OTHER INFORMATION: n = A,T,C or G  
US-10-087-192-97

Query Match 35.6%; Score 28.8; DB 13; Length 233060;  
Best Local Similarity 62.5%; Pred. No. 81;  
Matches 45; Conservative 0; Mismatches 27; Indels 0; Gaps 0;  
QY 5 ATTTGAAAAGATTGAGGCTGTTGTTGGATGTTGATGTTGATGTTGCTGACACTGACGG 64  
DB 102441 ATTTGTTTCCATCATGACGAGTTGATGTTGATGCCAGTTCTTGGTGACAGTGTAG 102500  
QY 65 CTGTGCCGCTC 76  
DB 102501 CTTAGCTGCTC 102512

RESULT 9  
US-10-087-192-1615  
;; Sequence 1615, Application US/10087192  
;; Publication No. US20020182586A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Morris, David W.  
;; APPLICANT: Engelhard, Eric K.  
;; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
;; TITLE OF INVENTION: CANCER  
;; FILE REFERENCE: 529452000122  
;; CURRENT APPLICATION NUMBER: US/10/087,192  
;; CURRENT FILING DATE: 2002-03-01  
;; PRIOR APPLICATION NUMBER: US 09/747,377  
;; PRIOR FILING DATE: 2000-12-22  
;; PRIOR APPLICATION NUMBER: US 09/798,586  
;; PRIOR FILING DATE: 2001-03-02  
;; NUMBER OF SEQ ID NOS: 2059  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 1615  
;; LENGTH: 72314  
;; TYPE: DNA  
;; ORGANISM: Mus musculus  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: (1)...(72314)  
;; OTHER INFORMATION: n = A,T,C or G  
US-10-087-192-1615

Query Match 34.6%; Score 28; DB 13; Length 72314;  
Best Local Similarity 63.2%; Pred. No. 1e+02;  
Matches 43; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
QY 2 TGAATTTGAAAAGATTGAGGCTGTTGTTGGATGTTGATGTTGATGTTGCTGACACTGA 61  
DB 32121 TTAATTTTAAAAATTTGTTTAAATTTATCTTATGTTGAGGGTATTGTTGCTTACATATA 32180  
QY 62 CGGCTGTG 69  
DB 32181 TGTCTGTG 32188

RESULT 10  
US-10-087-192-1342/c  
;; Sequence 1342, Application US/10087192  
;; Publication No. US20020182586A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Morris, David W.  
;; APPLICANT: Engelhard, Eric K.  
;; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
;; TITLE OF INVENTION: CANCER  
;; FILE REFERENCE: 529452000122  
;; CURRENT APPLICATION NUMBER: US/10/087,192  
;; CURRENT FILING DATE: 2002-03-01  
;; PRIOR APPLICATION NUMBER: US 09/747,377  
;; PRIOR FILING DATE: 2000-12-22  
;; PRIOR APPLICATION NUMBER: US 09/798,586  
;; PRIOR FILING DATE: 2001-03-02  
;; NUMBER OF SEQ ID NOS: 2059  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 1342  
;; LENGTH: 87731  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-087-192-1342

Query Match 34.6%; Score 28; DB 13; Length 87731;  
Best Local Similarity 63.2%; Pred. No. 1.1e+02;  
Matches 43; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
QY 5 ATTTGAAAAGATTGAGGCTGTTGTTGGATGTTGATGTTGATGTTGCTGACACTGACGG 64

Db 70914 ATATGAAGGGCTGAGGCTAGTGTCTTTTAAATAGATCAGTCCCACTGTGTCT 70855

Qy 65 CTGTGCG 72

Db 70854 CTGGGCTG 70847

## RESULT 11

US-10-425-115-162687/c  
; Sequence 162687, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 162687  
; LENGTH: 614  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_79947C.1  
US-10-425-115-162687

Query Match 34.1%; Score 27.6; DB 18; Length 614;  
Best Local Similarity 60.8%; Pred. No. 28;  
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 8 TGAAGAAGATTGAGGCTGTTTGTGATGTATGTGATTTGTGCTGACACTGACGGCTG 67

Db 189 TGATGATGGTGATCTCCACAGCGTGATGTCGATGGGTTTGTCTGAGGAAGAGGGCGG 130

Qy 68 TGCCGGCTCATGGC 81

Db 129 TGTAGGATCAGGCG 116

## RESULT 12

US-10-425-115-137066  
; Sequence 137066, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 137066  
; LENGTH: 1383  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_56480C.1  
US-10-425-115-137066

Query Match 34.1%; Score 27.6; DB 18; Length 1383;  
Best Local Similarity 78.6%; Pred. No. 37;  
Matches 33; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 5 ATTTGAAAGATTGAGGCTGTTTGTGATGTGTGATTTGGAT 46

Db 1290 ATATGACAGATGAAGACTGTGTTGTGATATATGTTGGAT 1331

## RESULT 13

US-10-425-115-83971/c  
; Sequence 83971, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 83971  
; LENGTH: 263  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_176594C.1  
US-10-425-115-83971

Query Match 33.8%; Score 27.4; DB 18; Length 263;  
Best Local Similarity 75.6%; Pred. No. 25;  
Matches 34; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 6 TTTGAAAAGATTGAGGCTGTTTGTGATGTATGTGATTTGT 50

Db 194 TTAGAATAGATTGAAGCTGCTGTGTGTGTGTTCTGTATGT 150

## RESULT 14

US-10-437-963-16223/c  
; Sequence 16223, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 16223  
; LENGTH: 648  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_21993C.1  
US-10-437-963-16223

Query Match 33.8%; Score 27.4; DB 17; Length 648;  
Best Local Similarity 62.3%; Pred. No. 33;  
Matches 43; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 ATGAATTTGAAAGATTGAGGCTGTTTGTGATGTATGTGATTTGTGCTGACACTG 60

Db 101 ATGTTTTTGTGAATAATTTTGTGAGGATGTTATGAATCTGTAATGTCATT 42

Qy 61 ACGCTGTG 69

Db 41 AGATCTGTG 33

```

RESULT 15
US-10-437-963-71944
; Sequence 71944, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 71944
; LENGTH: 1763
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72369C.1
US-10-437-963-71944

Query Match      33.8%; Score 27.4; DB 17; Length 1763;
Best Local Similarity 75.6%; Pred. No. 47;
Matches 34; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      4 AATTGAAAGATTGAGGCTGTTGTTGATGATGATGATGATTT 48
Db      1348 AAGATGAAAGATGATGATGATGATGATGATGATGATTT 1392

Search completed: November 11, 2004, 02:08:35
Job time : 55.7881 secs.

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 17:09:58 ; Search time 76.1638 Seconds  
(without alignments)  
6187.359 Million cell updates/sec

Title: US-09-909-207-1  
Perfect score: 663  
Sequence: 1 CAATCGTCACGACAAATTC.....TAACCTTCGATAAAACAAT 663

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:  
5: /cgn2\_6/prodata/1/ina/PCTUS\_COMB.seq:  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	663	100.0	663	3	US-08-470-953A-1
2	663	100.0	663	3	US-08-470-953A-2
3	663	100.0	744	3	US-08-470-953A-4
4	663	100.0	744	3	US-08-470-953A-5
5	663	100.0	1513	3	US-08-470-953A-10
6	663	100.0	1513	3	US-08-470-953A-11
7	634.2	95.7	744	3	US-09-189-060B-1
8	634.2	95.7	744	3	US-08-698-978-1
9	624.6	94.2	744	3	US-08-501-126-18
10	507.8	76.6	744	3	US-09-189-060B-11
11	255.4	38.5	2364	3	US-09-390-234-23
12	255.4	38.5	2364	4	US-09-603-311-23
13	255.2	38.5	1022	3	US-08-275-526C-1
14	255.2	38.5	1022	3	US-08-275-526C-35
15	255.2	38.5	1022	4	US-09-076-677-1
16	255.2	38.5	1022	4	US-09-076-677-35
17	255.2	38.5	1022	4	US-09-073-055-1
18	255.2	38.5	1022	4	US-09-073-055-35
19	255	38.5	600	3	US-08-275-526C-26
20	255	38.5	600	3	US-08-275-526C-34
21	255	38.5	600	4	US-09-076-677-26
22	255	38.5	600	4	US-09-076-677-34
23	255	38.5	600	4	US-09-073-055-26
24	255	38.5	600	4	US-09-073-055-34
25	255	38.5	681	3	US-08-275-526C-30
26	255	38.5	681	3	US-08-275-526C-32
27	255	38.5	681	4	US-09-076-677-30

28	255	38.5	681	4	US-09-076-677-32
29	255	38.5	681	4	US-09-073-055-30
30	255	38.5	681	4	US-09-073-055-32
31	204.8	30.9	1244	3	US-08-591-685-12
32	154.4	23.3	164	3	US-08-501-126-15
33	154.4	23.3	164	3	US-08-501-126-16
34	132.4	20.0	294	3	US-08-817-946-4
35	132.4	20.0	294	4	US-09-639-354A-4
36	119.4	18.0	1375	2	US-08-468-812-1
37	119.4	18.0	1375	3	US-08-590-563-1
38	119.4	18.0	1375	4	US-09-770-621-1
39	119.4	18.0	1375	4	US-09-235-832-1
40	119	17.9	229	3	US-09-189-060B-53
41	118.4	17.9	573	1	US-08-709-912-18
42	118.4	17.9	573	2	US-09-047-370-18
43	117.8	17.8	1207	1	US-08-575-964-2
44	117.8	17.8	1207	2	US-08-963-500-2
45	115.2	17.4	213	3	US-09-189-060B-49

ALIGNMENTS

RESULT 1  
US-08-470-953A-1  
Sequence 1, Appli US/08470953A  
Patent No. 6346407

GENERAL INFORMATION:

APPLICANT: ANDREE LAHAYE  
APPLICANT: ERIC DE BUYL  
APPLICANT: PIERRE LEDOUX  
APPLICANT: RENE DETROZ  
TITLE OF INVENTION: xylanase, microorganisms produced it,  
TITLE OF INVENTION: DNA molecule, processes for preparation of this xylanase  
TITLE OF INVENTION: and uses thereof  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:

ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
STREET: 2000 K St., N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/470,953A  
FILING DATE: 6-OCTOBER-1995  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilhem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-40  
TELEPHONE: 202-429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 663 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: Bacillus  
US-08-470-953A-1

Query Match 100.0%; Score 663; DB 3; Length 663;  
Best Local Similarity 100.0%; Pred. No. 7.9e-198;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAATCGTCACCGCAATTCATTTGGCAACACGATGCTATGATTAATTTTGAAA 60  
DB 1 CAATCGTCACCGCAATTCATTTGGCAACACGATGCTATGATTAATTTTGAAA 60

QY 61 GATAGCGTGGCTCTGGGACAATGATTTCTCAATCATGCGCGTACGTTGAGTCCCAATGG 120  
DB 61 GATAGCGTGGCTCTGGGACAATGATTTCTCAATCATGCGCGTACGTTGAGTCCCAATGG 120

QY 121 AACAAATGTTAAACAATATTTTCGTAAGGTAAATAATTCATGAACAACAACACAC 180  
DB 121 AACAAATGTTAAACAATATTTTCGTAAGGTAAATAATTCATGAACAACAACACAC 180

QY 181 CAACAAGTTGGTAACATGTCATTAACACTACGAGGCAACTTCCAAACCAATGGTAATGG 240  
DB 181 CAACAAGTTGGTAACATGTCATTAACACTACGAGGCAACTTCCAAACCAATGGTAATGG 240

QY 241 TATTTATGCGTCTATGTTGGACTGTTGACCTCTTGTGCAATATTAATTTGTCAGAGT 300  
DB 241 TATTTATGCGTCTATGTTGGACTGTTGACCTCTTGTGCAATATTAATTTGTCAGAGT 300

QY 301 TGGGCAACTGGCTCCACGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 360  
DB 301 TGGGCAACTGGCTCCACGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 360

QY 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAAGGGATGCCCACA 420  
DB 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAAGGGATGCCCACA 420

QY 421 TTTAAACAATATTTGAGTGTTCGAAGTCGAAACGACGAGTGGCAGCATTTCTGTCCAGC 480  
DB 421 TTTAAACAATATTTGAGTGTTCGAAGTCGAAACGACGAGTGGCAGCATTTCTGTCCAGC 480

QY 481 AACCACTTTAGAGCGTGGGAAACCTTAGGATGAATATGGGAAATGTATGAAGTCGCG 540  
DB 481 AACCACTTTAGAGCGTGGGAAACCTTAGGATGAATATGGGAAATGTATGAAGTCGCG 540

QY 541 CTCTACTGTAGAAGGCTATCAAAAGTAGCGGAAGTGTCTATATAGCAATACACTAAGA 600  
DB 541 CTCTACTGTAGAAGGCTATCAAAAGTAGCGGAAGTGTCTATATAGCAATACACTAAGA 600

QY 601 ATTAACGGTAACCTCTCTCAATATATAGTATGACGAGCATTAACCTTTGGATAAACA 660  
DB 601 ATTAACGGTAACCTCTCTCAATATATAGTATGACGAGCATTAACCTTTGGATAAACA 660

QY 661 AAT 663  
DB 661 AAT 663

RESULT 2  
US-08-470-953A-2  
; Sequence 2, Application US/08470953A  
; Patent No. 6346407  
; GENERAL INFORMATION:  
; APPLICANT: ANDREE LAHAYE  
; APPLICANT: ERIC DE BUIL  
; APPLICANT: PIERRE LEDOUX  
; APPLICANT: RENE DETROZ  
; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
; TITLE OF INVENTION: DNA molecule, processes for preparation of this xylanase  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
; STREET: 2000 K St., N.W., Suite 200  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.

APP!

ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/470,953A  
FILING DATE: 6-OCTOBER-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilhem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-40  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 663 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: Bacillus  
FEATURE:  
NAME/KEY: mat peptide  
LOCATION: 1..663  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..663  
US-08-470-953A-2

Query Match 100.0%; Score 663; DB 3; Length 663;  
Best Local Similarity 100.0%; Pred. No. 7.9e-198;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAATCGTCACCGCAATTCATTTGGCAACACGATGCTATGATTAATTTTGAAA 60  
DB 1 CAATCGTCACCGCAATTCATTTGGCAACACGATGCTATGATTAATTTTGAAA 60

QY 61 GATAGCGTGGCTCTGGGACAATGATTTCTCAATCATGCGCGTACGTTGAGTCCCAATGG 120  
DB 61 GATAGCGTGGCTCTGGGACAATGATTTCTCAATCATGCGCGTACGTTGAGTCCCAATGG 120

QY 121 AACAAATGTTAAACAATATTTTCGTAAGGTAAATAATTCATGAACAACAACACAC 180  
DB 121 AACAAATGTTAAACAATATTTTCGTAAGGTAAATAATTCATGAACAACAACACAC 180

QY 181 CAACAAGTTGGTAACATGTCATTAACACTACGAGGCAACTTCCAAACCAATGGTAATGG 240  
DB 181 CAACAAGTTGGTAACATGTCATTAACACTACGAGGCAACTTCCAAACCAATGGTAATGG 240

QY 241 TATTTATGCGTCTATGTTGGACTGTTGACCTCTTGTGCAATATTAATTTGTCAGAGT 300  
DB 241 TATTTATGCGTCTATGTTGGACTGTTGACCTCTTGTGCAATATTAATTTGTCAGAGT 300

QY 301 TGGGCAACTGGCTCCACGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 360  
DB 301 TGGGCAACTGGCTCCACGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 360

QY 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAAGGGATGCCCACA 420  
DB 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAAGGGATGCCCACA 420

QY 421 TTTAAACAATATTTGAGTGTTCGAAGTCGAAACGACGAGTGGCAGCATTTCTGTCCAGC 480  
DB 421 TTTAAACAATATTTGAGTGTTCGAAGTCGAAACGACGAGTGGCAGCATTTCTGTCCAGC 480



SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/470,953A

FILING DATE: 6-OCTOBER-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Wilhem F. Gadiano, Esq.

REGISTRATION NUMBER: 37,136

REFERENCE/DOCKET NUMBER: 4121-40

TELEPHONE: 202-429-0625

TELEFAX: (202) 293-1850

TELEX: 650 383-5605

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 744 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

STRAIN: Bacillus

FEATURE:

NAME/KEY: CDS

LOCATION: 1..744

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: 82..744

FEATURE:

NAME/KEY: sig\_peptide

LOCATION: 1..81

US-08-470-953A-5

Query Match 100.0%; Score 663; DB 3; Length 744;

Best Local Similarity 100.0%; Pred. No. 8.4e-198; Indels 0; Gaps 0;

Matches 663; Conservative 0; Mismatches 0;

QY 1 CAAATCGTCCACGACAAATCCATTTGGCAACACGATGGCTATGATTAATTTGGAAA 60

DB 82 CAAATCGTCCACGACAAATCCATTTGGCAACACGATGGCTATGATTAATTTGGAAA 141

QY 61 GATAGCGGTGGCTCTGGGCAATGATTTCAATCATCTGGGGTACCTTCAATGATTTGGAAA 120

DB 142 GATAGCGGTGGCTCTGGGCAATGATTTCAATCATCTGGGGTACCTTCAATGATTTGGAAA 201

QY 121 AACAAATGTTAAACAAATATTTATTTCCGTAAGGTAAATTTCAATGAAACACAAACACAC 180

DB 202 AACAAATGTTAAACAAATATTTATTTCCGTAAGGTAAATTTCAATGAAACACAAACACAC 261

QY 181 CAACAAATGTTAAACAAATGATTTCAATGATTTCAATGATTTCAATGATTTCAATGATTTGG 240

DB 262 CAACAAATGTTAAACAAATGATTTCAATGATTTCAATGATTTCAATGATTTCAATGATTTGG 321

QY 241 TATTTATGCTCTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGG 300

DB 322 TATTTATGCTCTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGG 381

QY 301 TGGGGCAACTGGGGTCCACGAGGACCAACGCTTAAGGGACCAATCTGTTGATGAGGA 360

DB 382 TGGGGCAACTGGGGTCCACGAGGACCAACGCTTAAGGGACCAATCTGTTGATGAGGA 441

QY 361 ACATATGATATCTACGAGATCTTAGATGATCAATCAACCTCCATTAAGGGATTTGCCACA 420

DB 442 ACATATGATATCTACGAGATCTTAGATGATCAATCAACCTCCATTAAGGGATTTGCCACA 501

QY 421 TTTTAAACAATTTGAGTGTTCGAGATCGAAACGACGAGTGGCAGATTTCTGTGAGC 480

DB 502 TTTTAAACAATTTGAGTGTTCGAGATCGAAACGACGAGTGGCAGATTTCTGTGAGC 561

QY 481 AACCACTTTAGAGCGTGGGAAAACCTTAGGATGATATGGGAAAATGATGATGATGATGATGATG 540

DB 562 AACCACTTTAGAGCGTGGGAAAACCTTAGGATGATATGGGAAAATGATGATGATGATGATGATG 621

QY 541 CTTACTGTAGAAGGCTATCAAAAGTAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA 600

DB 622 CTTACTGTAGAAGGCTATCAAAAGTAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA 681

QY 601 ATTAACGGTAAACCTCTCTCAACTATTAGTATACGAGAGCATAACTTTGGATAAAAAC 660

DB 682 ATTAACGGTAAACCTCTCTCAACTATTAGTATACGAGAGCATAACTTTGGATAAAAAC 741

QY 661 AAT 663

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744

DB 742 AAT 744

QY 742 AAT 744





ORIGINAL SOURCE: Van Beckhoven W.C., Rudolf F.  
ORGANISM: Bacillus agaradherens  
STRAIN: NCIMB 40482  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 82..744  
US-08-698-978-1

Query Match 95.7%; Score 634.2; DB 1; Length 871;  
Best Local Similarity 97.3%; Pred. No. 9.7e-189;  
Matches 645; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 CAAATCGTCACGACAAATTCATTTGGCAACACGATGGCTATGATTATGAATTTTGGAAA 60  
Db |||||||  
QY 82 CAAATCGTCACGACAAATTCATTTGGCAACACGATGGCTATGATTATGAATTTTGGAAA 141  
Db |||||||  
QY 61 GATAGCGGTGGCTCTGGGCAATGATTTCTCAATCATGCGGCTACGTTCAAGTCCCAATGG 120  
Db |||||||  
QY 142 GATAGCGGTGGCTCTGGGCAATGATTTCTCAATCATGCGGCTACGTTCAAGTCCCAATGG 201  
Db |||||||  
QY 121 AACATGTTTAAACACATATNTTCCGTAAGGTAAATAATTCATGAACACACACACAC 180  
Db |||||||  
QY 202 AACATGTTTAAACACATATNTTCCGTAAGGTAAATAATTCATGAACACACACACAC 261  
Db |||||||  
QY 181 CAACAAGTTGGTAACATGTCCTATAAATCAAGGACCACTTCCAAACCAATGTAATGG 240  
Db |||||||  
QY 262 CAACAAGTTGGTAACATGTCCTATAAATCAAGGACCACTTCCAAACCAATGTAATGG 321  
Db |||||||  
QY 241 TATTTATGCGTCTATGTTGGACTGTTGACCTCTTGTGCAATATATTTGTCGACAGT 300  
Db |||||||  
QY 322 TATTTATGCGTCTATGTTGGACTGTTGACCTCTTGTGCAATATATTTGTCGACAGT 381  
Db |||||||  
QY 301 TGGGGCAACTGGCGTCCACGAGGACCAACGCTTAAGGGACCACTGCTGTCATGGAGGA 360  
Db |||||||  
QY 382 TGGGGCAACTGGCGTCCACGAGGACCAACGCTTAAGGGACCACTGCTGTCATGGAGGA 441  
Db |||||||  
QY 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTTCCATTAAGGGATGGCCACA 420  
Db |||||||  
QY 442 ACATATGATATCTATGAACTCTTAGAGTCAATCAACCTTCCATTAAGGGATGGCCACA 501  
Db |||||||  
QY 421 TTTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACGAGTGGCACGATTTCTGTCAGC 480  
Db |||||||  
QY 502 TTTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACGAGTGGCACGATTTCTGTCAGC 561  
Db |||||||  
QY 481 AACCACTTTAGAGCGTGGGAAAACCTTAGGATGAATATGGGGAATATGATGAAGTCGG 540  
Db |||||||  
QY 562 AACCACTTTAGAGCGTGGGAAAACCTTAGGATGAATATGGGGAATATGATGAAGTCGG 621  
Db |||||||  
QY 541 CTTACTGTAGAAGCTATCAAGTACGAGGAGTGAATGATGATATAGCAATACACTAAGA 600  
Db |||||||  
QY 622 CTTACTGTAGAAGCTATCAAGTACGAGGAGTGAATGATGATATAGCAATACACTAAGA 681  
Db |||||||  
QY 601 ATTAACGGTAACCTCTCTCAACTATTTAGTAATGACGAGAGCACTAATTTGGATATAAAC 660  
Db |||||||  
QY 682 ATTAACGGTAACCTCTCTCAACTATTTAGTAATGACGAGAGCACTAATTTGGATATAAAC 741  
Db |||||||  
QY 661 AAT 663  
Db |||||  
QY 742 AAT 744  
Db |||||

APPLICANT: Van Beckhoven W.C., Rudolf F.  
APPLICANT: Quax, Wilhelmus J.  
APPLICANT: Jones, Brian E.  
TITLE OF INVENTION: ALKALI-TOLERANT XYLANASES  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morrison & Foerster  
STREET: 2000 Pennsylvania Avenue, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/501.126  
FILING DATE: 29-DEC-1995  
CLASSIFICATION: 425  
ATTORNEY/AGENT INFORMATION:  
NAME: Murahise, Kate H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 4615-0057.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 744 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
STRAIN: 1-43-3  
INDIVIDUAL ISOLATE: CBS672.93  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..744  
OTHER INFORMATION: /product= "xylanase"  
US-08-501-126-18

Query Match 94.2%; Score 624.6; DB 3; Length 744;  
Best Local Similarity 96.4%; Pred. No. 9.1e-186;  
Matches 639; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 CAAATCGTCACGACAAATTCATTTGGCAACACGATGGCTATGATTATGAATTTTGGAAA 60  
Db |||||||  
QY 79 CAAATCGTCACGACAAATTCATTTGGCAACACGATGGCTATGATTATGAATTTTGGAAA 138  
Db |||||||  
QY 61 GATAGCGGTGGCTCTGGGCAATGATTTCTCAATCATGCGGCTACGTTCAAGTCCCAATGG 120  
Db |||||||  
QY 139 GATAGCGGTGGCTCTGGGCAATGATTTCTCAATCATGCGGCTACGTTCAAGTCCCAATGG 198  
Db |||||||  
QY 121 AACATGTTTAAACACATATNTTCCGTAAGGTAAATAATTCATGAACACACACACAC 180  
Db |||||||  
QY 199 AATAATGTTTAAACATATATTTTCCGTAAGGTAAATAATTCATGAACACACACACAC 258  
Db |||||||  
QY 181 CAACAAGTTGGTAACATGTCCTATAAATCAAGGACCACTTCCAAACCAATGTAATGG 240  
Db |||||||  
QY 259 CAACAAGTTGGTAACATGTCCTATAAATCAAGGACCACTTCCAAACCAATGTAATGG 318  
Db |||||||  
QY 241 TATTTATGCGTCTATGTTGGACTGTTGACCTCTTGTGCAATATATTTGTCGACAGT 300  
Db |||||||  
QY 319 TATTTATGCGTCTATGTTGGACTGTTGACCTCTTGTGCAATATATTTGTCGACAGT 378  
Db |||||||  
QY 301 TGGGGCAACTGGCGTCCACGAGGACCAACGCTTAAGGGACCACTGCTGTCATGGAGGA 360  
Db |||||||  
QY 379 TGGGGCAACTGGCGTCCACGAGGACCAACGCTTAAGGGACCACTGCTGTCATGGAGGA 438  
Db |||||||

RESULT 9  
US-08-501-126-18  
Sequence 18, Application US/08501126  
Patent No. 6140095  
GENERAL INFORMATION:  
APPLICANT: Van Solingen, Pieter  
APPLICANT: Williams, Diane P.  
APPLICANT: Iverson, Sara  
APPLICANT: Farrell, Roberta L.  
APPLICANT: Herbes, Wilhelmina T.  
APPLICANT: Van Der Kleij, Wilhelmus A.  
APPLICANT: Herweijer, Margaretha A.



QY	361	ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCCA	420	Db	322	TATTTATGGCTCTATGGTTGGACTGTGACCTCTTGTGCAATATTTATATTGCGATAGT	381
Db	439	ACATATGATATCTATGAACCTCTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCCA	498	QY	301	TGGGGCAACTGGCGTCCACAGAGCAACGCTTAAGGGACCATCACTGTTGATGGAGA	360
QY	421	TTTAAACAATATTTGAGTGTTCGAAGATCGAAACGACGAGTGCACGATTTCTGTGAGC	480	Db	382	TGGGGCAACTGGCGTCCACAGAGCAACGCTTAAGGGACCATCACTGTTGATGGAGA	441
Db	499	TTTAAACAATATTTGAGTGTTCGAAGATCGAAACGACGAGTGCACGATTTCTGTGAGC	558	QY	361	ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCCA	420
QY	481	AACCACTTTAGAGCGTGGGAAACCTTAGGATGAATATGGGAAATGTATGAAGTGGCG	540	Db	442	ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCCA	501
Db	559	AACCACTTTAGAGCGTGGGAAACCTTAGGATGAATATGGGAAATGTATGAAGTGGCG	618	QY	421	TTTAAACAATATTTGAGTGTTCGAAGATCGAAACGACGAGTGCACGATTTCTGTGAGC	480
QY	541	CTTACTGTAGAGGCTATCAAGTAGCGAAGTGTATATATAGCAATACACTAAGA	600	Db	502	TTCAATCAGTACTGGTTCGATTCGACAGACGACGAGCGGCACTGTCACTACGGCA	561
Db	619	CTTACTGTAGAGGCTATCAAGTAGCGAAGTGTATATATAGCAATACACTAAGA	678	QY	481	AACCACTTTAGAGCGTGGGAAACCTTAGGATGAATATGGGAAATGTATGAAGTGGCG	540
QY	601	ATTAACGGTAAACCTCTCTCAACTATTAGTAAAGTAGAGTGTATATAGCAATACACTA	660	Db	562	AACCACTTTAATGCTGGGCTCTCTGGCATGAATATGGGTCATCAATACCAATC	621
Db	679	ATTAACGGTAAACCTCTCTCAACTATTAGTAAAGTAGAGTGTATATAGCAATACACTA	738	QY	541	CTTACTGTAGAGGCTATCAAGTAGCGAAGTGTATATATAGCAATACACTAAGA	600
QY	661	AAT 663		Db	622	CTGTTACTGAGGCTACCAATCTACGGAAGTCTAATGTATATAGCAATACACTAAGA	681
Db	739	AAT 741		QY	601	ATTAACGGTAAACCTCTCTCAACTATTAGTAAAGTAGAGTGTATATAGCAATACACTA	660
RESULT 10				Db	682	ATTAACGGTAAACCTCTCTCAACTATTAGTAAAGTAGAGTGTATATAGCAATACACTA	741
US-09-189-060B-11				QY	661	AAT 663	
; Sequence 11, Application US/09189060B				Db	742	AAT 744	
; Patent No. 6270968				RESULT 11			
; GENERAL INFORMATION:				US-09-390-234-23			
; APPLICANT: Dalboge, Henrik				; Sequence 23, Application US/09390234			
; APPLICANT: Sandal, Thomas				; Patent No. 6365390			
; APPLICANT: Kauppinen, Markus				; GENERAL INFORMATION:			
; APPLICANT: Borge, Diderichsen				; APPLICANT: Blum, David L.			
; TITLE OF INVENTION: Method Of Providing No. 6270968el DNA Sequences				; APPLICANT: Kataeva, Irina			
; FILE REFERENCE: 4772.204-US				; APPLICANT: Li, Xin-Liang			
; CURRENT APPLICATION NUMBER: US/09/189,060B				; APPLICANT: Ljungdahl, Lars G.			
; PRIOR FILING DATE: 1998-11-10				; TITLE OF INVENTION: Phenolic Acid Esterases, Coding Sequences and Methods			
; PRIOR APPLICATION NUMBER: PCT/DK97/00216				; FILE REFERENCE: 67-98			
; PRIOR FILING DATE: 1997-05-12				; CURRENT APPLICATION NUMBER: US/09/390,234			
; NUMBER OF SEQ ID NOS: 74				; CURRENT FILING DATE: 1999-09-03			
; SOFTWARE: FastSeq for Windows Version 4.0				; EARLIER APPLICATION NUMBER: US 60/099,136			
; SEQ ID NO 11				; EARLIER FILING DATE: 1998-09-04			
; TYPE: DNA				; NUMBER OF SEQ ID NOS: 24			
; ORGANISM: Hybrid				; SOFTWARE: Patent In Ver. 2.0			
; FEATURE:				; SEQ ID NO 23			
; NAME/KEY: CDS				; LENGTH: 2364			
; LOCATION: (1)...(744)				; TYPE: DNA			
US-09-189-060B-11				; ORGANISM: Clostridium stercoararium			
Query Match				; FEATURE:			
Best Local Similarity				; NAME/KEY: CDS			
Matches 566; Conservative				; LOCATION: (440)..(1975)			
QY	1	CAAAATCGTCCACCAATTCATTCGCAACCAACGATGGCTATGATTAATGAAATTTGGAAA	60	US-09-390-234-23			
Db	82	CAAAATCGTCCACCAATTCATTCGCAACCAACGATGGCTATGATTAATGAAATTTGGAAA	141	Query Match			
QY	61	GATAGCGGTGGCTCTGGGCAATGATTCATATCATGCGGCTACGTTAGTCCCAATGG	120	Best Local Similarity			
Db	142	GATAGCGGTGGCTCTGGGCAATGATTCATATCATGCGGCTACGTTAGTCCCAATGG	201	Matches 405; Conservative			
QY	121	AACAATGTTAAACATATTTCCGTAAGGTAAATTCATGAAATTCATGAAACACACACAC	180	QY	1	CAAAATCGTCCACCAATTCATTCGCAACCAACGATGGCTATGATTAATGAAATTTGGAAA	60
Db	202	AACAATGTTAAACATATTTCCGTAAGGTAAATTCATGAAATTCATGAAACACACACAC	261	Db	533	CGAATAATTTACGACAAATGAGACAGGCGCACATGGAGGCTACGATATGAGCTCTGGAAA	592
QY	181	CAACAAGTGGTAAACATGTCATATAACTACGAGGCAACCTTCCAAACCAATGTAATGCG	240	QY	61	GATAGCGGTGGCTCTGGGCAATGATTCATATCATGCGGCTACGTTAGTCCCAATGG	120
Db	262	CAACAAGTGGTAAACATGTCATATAACTACGAGGCAACCTTCCAAACCAATGTAATGCG	321	Db	593	GA---CTAGGAATACGATTTGGAATTAACGCGGTGGTACTTTTAGTTGCAATGG	649
QY	241	TATTTATGCGGTCTAAGTTGGACTGTGGACCTCTTGTGCAATATTATATTGTCGACAGT	300	QY	121	AACAATGTTAAACATATTTCCGTAAGGTAAATTCATGAAATTCATGAAACACACACAC	180
				Db	650	AGTAATATCGGTAAATGCACTATTAGAAAAGGGAGAAAATTTAATTCGACAAAACCTAT	709



QY 181 CAACAAGTTGGTAACATGTCCTCAATAAATACGAGGCAACTTCCACCAAAATGGTAATGG 240  
 Db 710 CAAGAATATTAGGAGACATAGTAGTGAATATGGCTGTGATTACAATCCAAACGGAATTTCC 769  
 QY 241 TATTATATGCTCTATGTTGGAGTGTGACCCCTTTGTGCAATATATTATTTGTCGACAGT 300  
 Db 770 TATTATGTTTACGGTTGGACAGAAATCCACTGGTTGAAATATTACATTTGTAGAAAGC 829  
 QY 301 TGGGGCAACTGGCGTCCACGAGGCAACGCTTAAGGGGACCAATCAC---TGTGATGGA 357  
 Db 830 TGGGGCAGCTGGCGTCCACTGGAGCAACACCCAAAGGAACCAATCACACAGTGGATGGCA 889  
 QY 358 GGAACATATGATATACAGAGCTCTTAGAGTCAATCAACCTTCCATTAAGGGGATGGC 417  
 Db 890 GGTACTTATGAATATATGAAATACCCCGGGTAAATCAGCCTTCCATGATGAACTGG 949  
 QY 418 ACATTTAAACAATATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGCAATTTCTGTC 477  
 Db 950 ACATTTCAACAATATTGGAGTGTTCGTACATCCAGAGAAACAAAGCGGAACAATATCTGTC 1009  
 QY 478 AGCAACCACTTTAGACGCTGGGAAACCTTAGGATGAATATGGGAAATATGATGAAGTTC 537  
 Db 1010 ACTGAACATTTTAAACAGTGGGAAAGAAATGGGCATCGGAATGGGTAAGATGTATGAAGTT 1069  
 QY 538 GCGCTTACTGTAGAGGCTCAAGTAGGGAAGTGTCTAATCTATATAGCAATACACTA 597  
 Db 1070 GCTCTTACCGTTGAAGTTATCAGAGCAGTGGGTACGCTAATGTATACAGAAATGAAATC 1129  
 QY 598 AGAATTAACGGTAACCC 614  
 Db 1130 AGAATAGGTGCAAAATCC 1146

RESULT 12

US-09-603-311-23  
 ; Sequence 23, Application US/09603311  
 ; Patent No. 6602700  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Li, Xin-Liang  
 ; APPLICANT: Ljungdahl, Lars G.  
 ; APPLICANT: Azain, Michael J.  
 ; APPLICANT: Davies, Edward T.  
 ; APPLICANT: Shah, Ashit K.  
 ; APPLICANT: Blum, David L.  
 ; APPLICANT: Kataeva, Irina  
 ; TITLE OF INVENTION: Phenolic Acid Esterases, Coding Sequences and Methods  
 ; FILE REFERENCE: 67-98A  
 ; CURRENT APPLICATION NUMBER: US/09/603,311  
 ; CURRENT FILING DATE: 2000-06-21  
 ; PRIOR APPLICATION NUMBER: US 60/099,136  
 ; PRIOR FILING DATE: 1998-09-04  
 ; PRIOR APPLICATION NUMBER: 09/390,324  
 ; PRIOR FILING DATE: 1999-09-03  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 23  
 ; LENGTH: 2364  
 ; TYPE: DNA  
 ; ORGANISM: Clostridium stercorarium  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (440)..(1975)  
 ; US-09-603-311-23

Query Match 38.5%; Score 255.4; DB 4; Length 2364;  
 Best Local Similarity 65.6%; Pred. No. 7.9e-70;  
 Matches 405; Conservative 0; Mismatches 206; Indels 6; Gaps 2;  
 QY 1 CAATCGTACCGACAAATTCATTTGGCAACCAACGATGGCTATGATTAATTTTGGAAA 60  
 Db 533 CGAATRAATTACGACAAATGAGACAGGCACACATGGAGGCTACGACTATGAGTCTGAAA 592

QY 61 GATACGGTGGCTCTGGGACAAATGATTTCTCAATCATGCGGTACGTTTCAGTGCCTCAATGG 120  
 Db 593 GA---CTACGGAATACGATTATGGAACCTTAACGACGGTGGTACTTTTAGTTGTCAATGG 649  
 QY 121 AACAAATGTAACACATATTTATTCGGTAAAGTAAATAATTTCAATGAACACACAAACACAC 180  
 Db 650 AGTAAATATCGGTAAATGCACTATTTAGAAAAGGGAGAAAATTTAAATTCGACAAAACCTAT 709  
 QY 181 CAACAAGTTGGTAACATGTCCTAACTACGAGGCAACTTCCAAACCAAAATGGTAATGG 240  
 Db 710 CAAGAATATTAGGAGACATAGTAGTGAATATGGCTGTGATTACAATCCAAACGGAATTTCC 769  
 QY 241 TATTATATGCTCTATGTTGGAGTGTTCGACCCCTTTGTGCAATATATTATTTGTCGACAGT 300  
 Db 770 TATTATGTTTACGGTTGGACAGAAATCCACTGGTTGAAATATTACATTTGTAGAAAGC 829  
 QY 301 TGGGGCAACTGGCGTCCACGAGGCAACGCTTAAGGGGACCAATCAC---TGTGATGGA 357  
 Db 830 TGGGGCAGCTGGCGTCCACTGGAGCAACACCCAAAGGAACCAATCACACAGTGGATGGCA 889  
 QY 358 GGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTTCCATTAAGGGGATGGC 417  
 Db 890 GGTACTTATGAATATATGAAATACCCCGGGTAAATCAGCCTTCCATCGATGGAATCTGTC 949  
 QY 418 ACATTTAAACAATATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGCAATTTCTGTC 477  
 Db 950 ACATTTCAACAATATTGGAGTGTTCGTACATCCAGAGAAACAAAGCGGAACAATATCTGTC 1009  
 QY 478 AGCAACCACTTTAGACGCTGGGAAACCTTAGGATGAATATGGGAAATATGATGAAGTTC 537  
 Db 1010 ACTGAACATTTTAAACAGTGGGAAAGAAATGGGCATCGGAATGGGTAAGATGTATGAAGTT 1069  
 QY 538 GCGCTTACTGTAGAGGCTCAAGTAGGGAAGTGTCTAATCTATATAGCAATACACTA 597  
 Db 1070 GCTCTTACCGTTGAAGTTATCAGAGCAGTGGGTACGCTAATGTATACAGAAATGAAATC 1129  
 QY 598 AGAATTAACGGTAACCC 614  
 Db 1130 AGAATAGGTGCAAAATCC 1146

RESULT 13

US-08-275-526C-1  
 ; Sequence 1, Application US/08275526C  
 ; Patent No. 6180382  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DE BUYL, ERIC  
 ; APPLICANT: LAHAYE, ANDR E  
 ; APPLICANT: LEDOUX, PIERRE  
 ; APPLICANT: AMORY, ANTOINE  
 ; APPLICANT: DETROZ, REN  
 ; APPLICANT: ANDRE, CHRISTOPHE  
 ; APPLICANT: VETTER, ROMAN  
 ; TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,  
 ; TITLE OF INVENTION: EXPRESSION VECTORS FOR SUCH XYLANASE AND  
 ; TITLE OF INVENTION: OTHER PROTEINS, HOST ORGANISMS THEREFOR AND  
 ; TITLE OF INVENTION: USE THEREOF  
 ; NUMBER OF SEQUENCES: 35  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.  
 ; STREET: 2000 K St., N.W., Suite 200  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20006  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/275,526C  
 ; FILING DATE: 15-JUL-1994

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gadiano, Wilhem F.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-49  
TELEPHONE: (202) 429-0625  
TELEFAX: (202) 293-0625  
TELEX: 650 383 5605  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1022 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
ORIGINAL SOURCE: Bacillus pumilus  
ORGANISM: PRL B12  
STRAIN: PRL B12  
US-08-275-526C-1

Query Match 38.5%; Score 255.2; DB 3; Length 1022;  
Best Local Similarity 65.5%; Pred. No. 5.9e-70;  
Matches 406; Conservative 0; Mismatches 208; Indels 6; Gaps 2;

QY 13 GACAAATCCATTGGCAACACGATGGCTATGATTGAAATTTTGGAAAGATAGCGGTGGC 72  
DB 279 GATAATAGGATAGGACACACAGCGGATACGATTTTGAATTTATGGAAGATTAC---GGA 335

QY 73 TCTGGGCAATGATCTCAATCATGCGGTACGTTTCAGTCCCAATGGAACAATGTTTAA 132  
DB 336 AATACCTCGATGACATCAATAACGCGGGGCACTTTAGTGAAGCTGGAACAATATTGGA 395

QY 133 AACATATTATTTCCGTAAGGTAAATAATTCATGAAACACAAACACACCAAGTTGGT 192  
DB 396 AATGCTTTATTTCCAAAGGAAAGAGTTTGAATTCCTAACTCACTCACTTGGC 455

QY 193 AACATGTCATPAACTAGCGAGCCAACTTCACCAAAATGTTATGTTATGTTATGTC 252  
DB 456 AACATCTCCATCAACTACACGCGACCTTTAAACCGGGCGGAAATTCCTATTTATGTC 515

QY 253 TATGTTGGACTGTTGACCTCTTGTGCAATATTATATTTCGACAGTTGGGCAACTGG 312  
DB 516 TATGGCTGACACATCTCATTTAGCTGAATACATTTGTTAGTCAATGAGGACATAT 575

QY 313 CGTCCACGAGGACAGCCTAAGGGACCATCATCTGTTGATGGAGGACATATGATATC 372  
DB 576 CGTCCACAGG---AACGTATAAAGGATCAATTTTATGCGGATGGAGGACATATGACATA 632

QY 373 TAGGACTCTTAGAGTCAATCAACCTCATTAAGGGGATTCGACATTTTAAACAATAT 432  
DB 633 TATGAAACGCTCGGTGTCATACAGCTTCTATCATTTGGAGAGCTACCTTCAACAATAT 692

QY 433 TGGAGTGTTCGAAGATCGAAACGACAGTGGCAGCAATTTCTGTCAGCAACCACTTTAGA 492  
DB 693 TGGAGTGTAGTCAACAAACAAACGACAAAGCGGAAACGCTCTCGGTGAGTGGAGCTTTAAA 752

QY 493 GCGTGGGAAACTTAGGATGAATATGCGGAAATGATGAAGTCGCGCTTACTGTAGAA 552  
DB 753 AAATGGGAAAGCTTAGCGATGCCAATGGGAAATATGTAAGAACAGCATTAATCTAGH 812

QY 553 GGCTATCAAAAGTAGCGAGTGTCTAATGTATATAGCAATACACTAAGAAATTAACGGTAA 612  
DB 813 GGCTACGAGACGAGGAGTGCATGTGTCACGAATCAGCTGATGATTCGATAAAG 872

QY 613 CCTCTCTCAACTATTAGTAA 632  
DB 873 CATATGAAAGGAGGACGAA 892

RESULT 14  
US-08-275-526C-35  
; Sequence 35, Application US/08275526C

Patent No. 6180382  
GENERAL INFORMATION:  
APPLICANT: DE BOYL, ERIC  
APPLICANT: LAHAYE, ANDR E  
APPLICANT: LEDOUX, PIERRE  
APPLICANT: AMORY, ANTOINE  
APPLICANT: DETROZ, REN  
APPLICANT: ANDRE, CHRISTOPHE  
APPLICANT: VETTER, ROMAN  
TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,  
TITLE OF INVENTION: EXPRESSION VECTORS FOR SUCH XYLANASE AND  
TITLE OF INVENTION: OTHER PROTEINS, HOST ORGANISMS THEREFOR AND  
TITLE OF INVENTION: USE THEREOF  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAN BRINKS HOFER GILSON & LIONE, P.C.  
STREET: 2000 K St., N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/275,526C  
FILING DATE: 15-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gadiano, Wilhem F.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-49  
TELEPHONE: (202) 429-0625  
TELEFAX: (202) 293-0625  
TELEX: 650 383 5605  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1022 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
US-08-275-526C-35

Query Match 38.5%; Score 255.2; DB 3; Length 1022;  
Best Local Similarity 65.5%; Pred. No. 5.9e-70;  
Matches 406; Conservative 0; Mismatches 208; Indels 6; Gaps 2;

QY 13 GACAAATCCATTGGCAACACGATGGCTATGATTGAAATTTTGGAAAGATAGCGGTGGC 72  
DB 279 GATAATAGGATAGGACACACAGCGGATACGATTTTGAATTTATGGAAGATTAC---GGA 335

QY 73 TCTGGGCAATGATCTCAATCATGCGGTACGTTTCAGTCCCAATGGAACAATGTTTAA 132  
DB 336 AATACCTCGATGACATCAATAACGCGGGGCACTTTAGTGAAGCTGGAACAATATTGGA 395

QY 133 AACATATTATTTCCGTAAGGTAAATAATTCATGAAACACAAACACACCAAGTTGGT 192  
DB 396 AATGCTTTATTTCCAAAGGAAAGAGTTTGAATTCCTAACTCACTCACTTGGC 455

QY 193 AACATGTCATPAACTAGCGAGCCAACTTCACCAAAATGTTATGTTATGTTATGTC 252  
DB 456 AACATCTCCATCAACTACACGCGACCTTTAAACCGGGCGGAAATTCCTATTTATGTC 515

QY 253 TATGTTGGACTGTTGACCTCTTGTGCAATATTATATTTCGACAGTTGGGCAACTGG 312  
DB 516 TATGGCTGACACATCTCATTTAGCTGAATACATTTGTTAGTCAATGAGGACATAT 575

QY 313 CGTCCACGAGGACAGCCTAAGGGACCATCATCTGTTGATGGAGGACATATGATATC 372  
DB 576 CGTCCACAGG---AACGTATAAAGGATCAATTTTATGCGGATGGAGGACATATGACATA 632

QY 373 TAGGACTCTTAGAGTCAATCAACCTCATTAAGGGGATTCGACATTTTAAACAATAT 432  
DB 633 TATGAAACGCTCGGTGTCATACAGCTTCTATCATTTGGAGAGCTACCTTCAACAATAT 692

QY 433 TGGAGTGTTCGAAGATCGAAACGACAGTGGCAGCAATTTCTGTCAGCAACCACTTTAGA 492  
DB 693 TGGAGTGTAGTCAACAAACAAACGACAAAGCGGAAACGCTCTCGGTGAGTGGAGCTTTAAA 752

QY 493 GCGTGGGAAACTTAGGATGAATATGCGGAAATGATGAAGTCGCGCTTACTGTAGAA 552  
DB 753 AAATGGGAAAGCTTAGCGATGCCAATGGGAAATATGTAAGAACAGCATTAATCTAGH 812

QY 553 GGCTATCAAAAGTAGCGAGTGTCTAATGTATATAGCAATACACTAAGAAATTAACGGTAA 612  
DB 813 GGCTACGAGACGAGGAGTGCATGTGTCACGAATCAGCTGATGATTCGATAAAG 872

QY 613 CCTCTCTCAACTATTAGTAA 632  
DB 873 CATATGAAAGGAGGACGAA 892

Db 576 CGTCCACAGG--AACGTATTAAGGATCATTTTATGCGGATGGAGGCACATATGACATA 632  
Qy 373 TAGGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATGGCCACATTTAAACAATAT 432  
Db 633 TATGAACCGCTCGGTCAATCAGCCTTCTATCATTTGGAGAGCGCTACCTTCAACAATAT 692  
Qy 433 TGGAGTGTTCGAAGATCGAAACGCGAGTGGCAGCATTTCTGTCAAGCAACCACTTTAGA 492  
Db 693 TGGAGTGTACGTCAAAACAAACGCAACGCGACGCTCTCGTCAAGGATTTTAA 752  
Qy 493 GCGTGGGAAACCTTAGGATGAATATGGGAAATGTATGAAGTCCGCTTACTGTAGAA 552  
Db 753 AAATGGGAAAGCTTAGCGATGCCAATGGGAAATGTATGAACACGACATTAACGTAGAA 812  
Qy 553 GCGTATCAAGTAGCGGAGTGTATGTATATAGCAATACACTAAGATTAACGGTAA 612  
Db 813 GCGTACCGAAGCAACGGAAGTGGCAATGTATGACGAATCAGCTGATGATTCGATAAAG 872  
Qy 613 CCTCTCTCAACTATTAGTAA 632  
Db 873 CATATGAAAAGCCAGCAA 892

RESULT 15

US-09-076-677-1  
Sequence 1, Application US/09076677  
Patent No. 6423523

GENERAL INFORMATION:

APPLICANT: DE BUYL, ERIC  
LAHAYE, ANDREE  
LEDOUX, PIERRE  
AMORY, ANTOINE  
DETROZ, RENE  
ANDRE, CHRISTOPHE  
VETTER, ROMAN

TITLE OF INVENTION:

XYLANASE DERIVED FROM A BACILLUS SPECIES,  
EXPRESSION VECTORS FOR SUCH XYLANASE AND  
OTHER PROTEINS, HOST ORGANISMS THEREFOR AND  
USE THEREOF

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P. C.  
STREET: 2000 K St., N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/076,677  
FILING DATE: 12-May-1998  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/275,526  
FILING DATE: 15-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Gadiano, Wilhem P.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-49  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 429-0625  
TELEFAX: (202) 293-0625  
TELEX: 650 383 5605

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1022 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: genomic DNA  
ORIGINAL SOURCE:  
ORGANISM: Bacillus pumilus  
STRAIN: PRL B12  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-076-677-1  
Query Match 38.5%; Score 255.2; DB 4; Length 1022;  
Best Local Similarity 65.5%; Pred. No. 5,9e-70;  
Matches 406; Conservative 0; Mismatches 208; Indels 6; Gaps 2;  
Qy 13 GACAATTCATTTGGCAACACGATGGCTATGATTTATGATTTTGGAAAGATAGCGGTGGC 72  
Db 279 GATAATAGGATAGGAGCACACAGCGGATACGATTTTGAATTTATGGAAGATTAC---GGA 335  
Qy 73 TCTGGGCAATGATTTCTCAATCATCGCGTACGTTTCAGTGGCCCAATGGAACCAATGTTAAC 132  
Db 336 AATACCTCGATGACACTCAATTAACCGCGGGGCAATTTAGTGGCAAGCTGGAAACAATATGGA 395  
Qy 133 AACATATTATTCGGTAAAGGTAAAAAATTTCAATGAACAACAACAACCAACCAAGTTGGT 192  
Db 396 AATGCTTTATTTGAAAAGGAAGAGTTTGTATCCACTAAATCTCATCACTTGGC 455  
Qy 193 AACATGTCCATAAACTACGGAGCCCACTTCCAAACCAATGGTAAATCGGTATTTATCGCTC 252  
Db 456 AACATCTCCATCAACTACAAACGAGCCTTTAAACCGCGGGGAAATTCCTATTTATGTGTC 515  
Qy 253 TATGGTTGGACTGTTGACCCCTCTTGTGCAATATATTTATGTCGACAGTTGGGSCAACTGG 312  
Db 516 TATGGCTGGACACAATCTCCATTTAGTGAATCTACTATTTGTTGATGTCATGGGCAATAT 575  
Qy 313 CGTCCACAGGAGCAACGCTTAAGGGGACCATCTGTTGATGGAGGAACATATGATATC 372  
Db 576 CGTCCACAGG---AACGTATTAAGGATCATTTTATGCGGATGGAGGCACATATGACATA 632  
Qy 373 TAGGAGACTCTTAGAGTCAATCAACCTCCATTTAAGGGGATTTGCCACATTTAAACAATAT 432  
Db 633 TATGAACCGCTCGGTCAATCAGCCTTCTATCATTTGGAGAGCGCTACCTTCAACAATAT 692  
Qy 433 TGGAGTGTTCGAAGATCGAAACGCGAGTGGCAGCATTTCTGTCAAGCAACCACTTTAGA 492  
Db 693 TGGAGTGTACGTCAAAACAAACGCAACGCGGACCGTCTCCGTCAAGTGGCAATTTAA 752  
Qy 493 GCGTGGGAAACCTTAGGGATGAATATGGGAAATGTATGAAGTCCGCTTACTGTAGAA 552  
Db 753 AAATGGGAAAGCTTAGGCATGCCAATGGGAAATGTATGAACACGACATTTAAGTAA 812  
Qy 553 GCGTATCAAGTAGCGGAGTGTCTAATGTATATAGCAATACACTAAGATTAACGGTAA 612  
Db 813 GCGTACCGAAGCAACGGAAGTGGCAATGTATGACGAATCAGCTGATGATTCGATAAAG 872  
Qy 613 CCTCTCTCAACTATTAGTAA 632  
Db 873 CATATGAAAAGCCAGCAA 892

Search completed: November 11, 2004, 01:34:19  
Job time : 79.1638 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 10, 2004, 15:55:17 ; Search time 420.566 Seconds  
(without alignments)  
8275.436 Million cell updates/sec

Title: US-09-909-207-1  
Perfect score: 663  
Sequence: 1 CAATCGTCACGACAATTC.....TAACTTTGGATAAACAAT 663

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 23Sep04: \*  
1: Geneseqn1980s: \*  
2: Geneseqn1990s: \*  
3: Geneseqn2000s: \*  
4: Geneseqn2001as: \*  
5: Geneseqn2001bs: \*  
6: Geneseqn2002as: \*  
7: Geneseqn2002bs: \*  
8: Geneseqn2003as: \*  
9: Geneseqn2003bs: \*  
10: Geneseqn2003cs: \*  
11: Geneseqn2003ds: \*  
12: Geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	663	100.0	663	2	AAT16101 Xylanase
2	663	100.0	744	2	AAT16102 Xylanase
3	663	100.0	1513	2	AAT16103 Xylanase
4	634.2	95.7	744	2	AAT13067 Glycosyl
5	634.2	95.7	871	2	AAT30255 DNA encod
6	624.6	94.2	744	2	AAT92878 Thermosta
7	507.8	76.6	747	2	AAT13074 Xylanase
8	371.2	56.0	1068	12	ADJ34947 DNA encod
9	276.4	41.7	1956	12	ADJ35051 DNA encod
10	255.4	38.5	2364	3	AAT21821 Clostridi
11	255.2	38.5	1022	2	AAT90923 B. pumilu
12	255	38.5	684	12	ADJ34969 DNA encod
13	212.2	32.0	1190	2	AAT90972 Nucleotid
14	204.8	30.9	1244	2	AAT08142 Xylanase
15	177.8	26.8	747	12	ADJ35011 DNA encod
16	154.4	23.3	164	2	AAT92876 Thermosta
17	151.2	22.8	164	2	AAT92875 Thermosta
18	143.6	21.7	1695	12	ADJ35101 DNA encod
19	141.2	21.3	1338	12	ADJ35151 DNA encod
20	136.2	20.5	1077	12	ADJ34965 DNA encod
21	130.2	19.6	1065	12	ADJ34999 DNA encod

22	128.4	19.4	1047	12	ADJ34949
23	121.2	18.3	678	12	ADJ34955
24	119.4	18.0	1375	2	AAX90405 Actinomad
25	119.4	18.0	1375	2	AAT64930 Actinomad
26	119	17.9	229	2	AAT15063 Xylanase
27	118.4	17.9	573	2	AAT36098 DNA seque
28	117.8	17.8	1207	2	AAT42374 Bacillus
29	117.6	17.7	1041	12	ADJ34943 DNA encod
30	116.8	17.6	1047	12	ADJ35083 DNA encod
31	115.2	17.4	213	2	AAT15059 Xylanase
32	115.2	17.4	596	3	AAT48219 T. reesei
33	115.2	17.4	596	6	AAT29410 Trichoder
34	115.2	17.4	596	9	AAT60925 Trichoder
35	115.2	17.4	596	9	AAT60941 Trichoder
36	112	16.9	1074	12	ADJ34983 DNA encod
37	112	16.9	1137	12	ADJ35039 DNA encod
38	111.8	16.9	1273	2	AAT90388 Xylanase
39	111.4	16.8	636	12	ADJ34981 DNA encod
40	111.4	16.8	942	10	AAT80366 A. fumiga
41	111.2	16.8	1008	12	ADJ35015 DNA encod
42	110.8	16.7	669	12	ADJ34951 DNA encod
43	110.4	16.7	1047	12	ADJ35037 DNA encod
44	110	16.6	1002	10	AAT80365 A. fumiga
45	109.4	16.5	1041	12	ADJ34953 DNA encod

ALIGNMENTS

RESULT 1  
AAT16101  
ID AAT16101 standard; DNA; 663 BP.  
XX AC AAT16101;  
XX 16-OCT-2003 (revised)  
DT 15-MAY-1996 (first entry)  
XX Xylanase gene.  
XX Xylanase; thermostable enzyme; paper; pulp; bleaching; feedstuff; baking;  
KW ss.  
XX Bacillus sp; strain 710/1 (LMG P-14798):  
XX Key Location/Qualifiers  
FT mat\_peptide 1..663  
FT /\*tag= a  
FT /EC\_number= "3.2.1.8"  
XX  
XX AU9525086-A.  
XX 08-FEB-1996.  
XX 19-JUL-1995; 95AU-00025086.  
XX 26-JUL-1994; 94BE-00000706.  
XX 17-MAY-1995; 95BE-00000448.  
XX (SOLV ) SOLVAY SA.  
XX De Buyl E, Labaye A, Ledoux P, Detroz R;  
XX WPI; 1996-117341/13.  
XX P-PSDB; AAR92053.  
XX Bacillus derived xylanase active over wide pH range - used in treatment  
XX of paper pulp, animal feeds and in bakery goods.  
XX Claim 30; Page 50-51; 94pp; English.  
XX A DNA sequence (AAT16101) coding for a thermostable mature xylanase  
XX (AAR92053) was isolated from a gene library of Bacillus sp. 720/1 (LMG P-  
CC

Applicant's  
work  
priority claimed

CC 14798). Sequences were also obtd. (see AAT16102 and AAT16103) for the  
CC xylanase precursor and for the complete gene including 5' and 3',  
CC untranslated sequences. DNA coding for the mature enzyme may be  
CC incorporated into a vector and expressed from either its own promoter or  
CC from the Bacillus pumilus PRL B12 promoter (AAQ73996), and used for  
CC prodn. of recombinant xylanase in transformed hosts, pref. Bacillus  
CC licheniformis or B. pumilus. The enzyme is useful in the paper-pulp,  
CC animal feed and baking industries. (Updated on 16-OCT-2003 to standardise  
CC OS field)

XX SQ Sequence 663 BP; 220 A; 131 C; 146 G; 166 T; 0 U; 0 Other;

Query Match 100.0%; Score 663; DB 2; Length 744;  
Best Local Similarity 100.0%; Pred. No. 5e-188;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAATCGTCACGCAATTCATGGCAACACGATGGCTATGATTAATGATTTGGAAA 60  
DB 1 CAATCGTCACGCAATTCATGGCAACACGATGGCTATGATTAATGATTTGGAAA 60  
QY 61 GATAGCGTGGCTCTGGGACAAATGATTCATCATGGCGGTACGTTCAAGTCCCAATGG 120  
DB 61 GATAGCGTGGCTCTGGGACAAATGATTCATCATGGCGGTACGTTCAAGTCCCAATGG 120  
QY 121 AACAAATGTTAAACAATATTTTCGTAAGGTAAAAATTCATTAAGAAACAACAACAC 180  
DB 121 AACAAATGTTAAACAATATTTTCGTAAGGTAAAAATTCATTAAGAAACAACAACAC 180  
QY 181 CAACAAGTGGTAACATGTCATTAACATGACGAGCAACCTCCACCAAAATGGTAATGCG 240  
DB 181 CAACAAGTGGTAACATGTCATTAACATGACGAGCAACCTCCACCAAAATGGTAATGCG 240  
QY 241 TATTATGCGTCTATGGTGGACTGTGACCCCTCTGTGCGAATATTAATTTGCGACAGT 300  
DB 241 TATTATGCGTCTATGGTGGACTGTGACCCCTCTGTGCGAATATTAATTTGCGACAGT 300  
QY 301 TGGGCAACTGGGCTCCACGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGA 360  
DB 301 TGGGCAACTGGGCTCCACGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGA 360  
QY 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATGCCACA 420  
DB 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATGCCACA 420  
QY 421 TTTAAACAATATGGAGTGTTCGAGATCGAAACGACGAGTGGCAGATTTCTGTGAGC 480  
DB 421 TTTAAACAATATGGAGTGTTCGAGATCGAAACGACGAGTGGCAGATTTCTGTGAGC 480  
QY 481 AACCCTTTAGAGCGTGGGAAACTTAGGATGAATATGGGAAATATGATGAATGCGG 540  
DB 481 AACCCTTTAGAGCGTGGGAAACTTAGGATGAATATGGGAAATATGATGAATGCGG 540  
QY 541 CTTACTGTAGAGGCTATCAAGTACGGAAGTCTTAATGTATATAGCAATACATAAGA 600  
DB 541 CTTACTGTAGAGGCTATCAAGTACGGAAGTCTTAATGTATATAGCAATACATAAGA 600  
QY 601 ATTAACGGTAACCTCTCTCAACTATAGTATGACGAGCATTAATTTGATATAAAC 660  
DB 601 ATTAACGGTAACCTCTCTCAACTATAGTATGACGAGCATTAATTTGATATAAAC 660  
QY 661 AAT 663  
DB 661 AAT 663

RESULT 2  
AAT16102  
ID AAT16102 standard; DNA; 744 BP.  
XX  
AC AAT16102;  
XX  
DT 16-OCT-2003 (revised)  
DT 15-MAY-1996 (first entry)

XX Xylanase precursor gene.  
DE Xylanase; thermostable enzyme; paper; pulp; bleaching; feedstuff; baking;  
KW ss.  
KW Bacillus sp; strain 710/1 (LMG P-14798).  
OS

XX Key Location/Qualifiers  
FH sig\_peptide 1..81 /tag= a  
FT mat\_peptide 82..744 /tag= b  
FT /EC\_number= "3.2.1.8"

XX AU9525086-A.  
XX 08-FEB-1996.  
XX 19-JUL-1995; 95AU-00025086.  
XX 26-JUL-1994; 94BE-00000706.  
XX 17-MAY-1995; 95BE-00000448.  
XX (SOLV ) SOLVAY SA.

XX De Buyl E, Lahaye A, Ledoux P, Detroz R;  
XX WPI; 1996-117341/13.  
XX P-PSDB; AAR92054.

XX Bacillus derived xylanase active over wide pH range - used in treatment  
XX of paper pulp, animal feeds and in bakery goods.  
XX Claim 9; Page 54-55; 94pp; English.

XX A DNA sequence (AAT16102) coding for a xylanase precursor (AAR92054) was  
XX isolated from a gene library of Bacillus sp. 720/1 (LMG P-14798). The  
XX sequence (AAT16103) for the complete gene including 5' and 3',  
XX untranslated sequences was also obtd. The gene may be incorporated into a  
XX vector and expressed from either its own promoter or from the Bacillus  
XX pumilus PRL B12 promoter (see AAQ73996), and used for prodn. of  
XX recombinant thermostable xylanase in transformed hosts, pref. Bacillus  
XX licheniformis or B. pumilus. The enzyme is useful in the paper-pulp,  
XX animal feed and baking industries. (Updated on 16-OCT-2003 to standardise  
XX OS field)

XX SQ Sequence 744 BP; 246 A; 144 C; 160 G; 194 T; 0 U; 0 Other;

Query Match 100.0%; Score 663; DB 2; Length 744;  
Best Local Similarity 100.0%; Pred. No. 5.3e-188;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAATCGTCACGCAATTCATGGCAACACGATGGCTATGATTAATGATTTGGAAA 60  
DB 82 CAATCGTCACGCAATTCATGGCAACACGATGGCTATGATTAATGATTTGGAAA 141  
QY 61 GATAGCGTGGCTCTGGGACAAATGATTCATCATGGCGGTACGTTCAAGTCCCAATGG 120  
DB 142 GATAGCGTGGCTCTGGGACAAATGATTCATCATGGCGGTACGTTCAAGTCCCAATGG 201  
QY 121 AACAAATGTTAAACAATATTTTCGTAAGGTAAAAATTCATTAAGAAACAACAACAC 180  
DB 202 AACAAATGTTAAACAATATTTTCGTAAGGTAAAAATTCATTAAGAAACAACAACAC 261  
QY 181 CAACAAGTGGTAACATGTCATTAACATGACGAGCAACCTCCACCAAAATGGTAATGCG 240.  
DB 262 CAACAAGTGGTAACATGTCATTAACATGACGAGCAACCTCCACCAAAATGGTAATGCG 321  
QY 241 TATTATGCGTCTATGGTGGACTGTGACCCCTCTGTGCGAATATTAATTTGCGACAGT 300  
DB 322 TATTATGCGTCTATGGTGGACTGTGACCCCTCTGTGCGAATATTAATTTGCGACAGT 381



AC AAV13067;  
 XX  
 DT 19-MAY-1998 (first entry)  
 XX  
 DE Glycosyl hydrolase family 11 xylanase DNA derived from *Bacillus* sp.  
 XX  
 KW *Bacillus* sp.; xylanase; glycosyl hydrolase family 11; isolation;  
 XX microorganism; identification; ss.  
 XX  
 OS *Bacillus* sp.  
 XX  
 FH Location/Qualifiers  
 FT 1. 747  
 FT /tag= a  
 FT /product= "glycosyl hydrolase family 11 xylanase"  
 FT  
 XX WO9743409-A2.  
 XX  
 XX 20-NOV-1997.  
 XX  
 XX 12-MAY-1997; 97WO-DK000216.  
 XX  
 XX 10-MAY-1996; 96DK-00000562.  
 XX  
 XX (NOVO ) NOVO-NORDISK AS.  
 XX  
 XX Dalboge H, Diderichsen B, Sandal T, Kauppinen S;  
 XX  
 XX WPI; 1998-008878/01.  
 DR P-PSDB; AAW44262.  
 DR  
 XX  
 PT Isolating novel DNA sequences from microorganisms - without the need for  
 PT culturing the microorganism.  
 PS  
 PS Example 1; Page 31-32; 72pp; English.  
 XX  
 CC The present sequence encodes a polypeptide with xylanase activity used in  
 CC an example of the present invention. The present invention describes a  
 CC novel method for providing a novel DNA sequence encoding a polypeptide  
 CC from a microorganism with an activity of interest. The method comprises:  
 CC (i) PCR amplification of the DNA with PCR primers with homology to (a)  
 CC known gene(s) encoding a polypeptide with an activity of interest; (ii)  
 CC linking the obtained PCR product of a 5' structural gene sequence and a  
 CC 3' structural gene sequence; (iii) expressing the resulting hybrid DNA  
 CC sequence; (iv) screening for hybrid DNA sequences encoding a polypeptide  
 CC with the activity of interest or a related activity; and (v) isolating  
 CC the hybrid DNA sequence identified in step (iv). This method provides for  
 CC identification and isolation of sequences from microorganisms without  
 CC having to cultivate and isolate the microorganism  
 SQ  
 SQ Sequence 744 BP; 249 A; 146 C; 157 G; 192 T; 0 U; 0 Other;  
 Query Match 95.7%; Score 634.2; DB 2; Length 744;  
 Best Local Similarity 97.3%; Pred. No. 2.2e-179;  
 Matches 645; Conservative 0; Mismatches 18; Indels 0; Gaps 0;  
 QY 1 CAAATCGTCACGCAATTCATTCGACACGATGGCTATGATTGAATTTGGAAA 60  
 DB 82 CAAATCGTCACGCAATTCATTCGACACGATGGCTATGATTGAATTTGGAAA 141  
 QY 61 GATAGCGGTGGCTCTGGGACATGATTTCTCAATCATGCGGTACGTTTCAGTCCCAATGG 120  
 DB 142 GATAGCGGTGGCTCTGGGACATGATTTCTCAATCATGCGGTACGTTTCAGTCCCAATGG 201  
 QY 121 AACATGTTTAAACAATATTTTCGTTAAAGTAAATAATTCATGAAGAACACACACAC 180  
 DB 202 AACATGTTTAAACAATATTTTCGTTAAAGTAAATAATTCATGAAGAACACACACAC 261  
 QY 181 CAACAAGTTGGTAAACATGTCATAAATACGAGGACCACTTCCAAACCAATGGTAATGG 240  
 DB 262 CAACAAGTTGGTAAACATGTCATAAATACGAGGACCACTTCCAAACCAATGGTAATGG 321  
 QY 241 TATTTATGCGTCTATGGTTGGACTGTGTGACCTCTGTGCAATATATATTTGTCACAGT 300

DB 322 TATTTATGCGTCTATGGTTGGAAGCTTGTGACCTCTTGTGCAATATATATTTGTCAGTAGT 381  
 QY 301 TGGGGCAACTGGCGTCCACCAGAGCAACCGCTTAAGGGGACCACCTACCTGTTGATGGAGGA 360  
 DB 382 TGGGGCAACTGGCGTCCACCAGGGGCAACCGCTTAAGGGGACCACCTACCTGTTGATGGAGGA 441  
 QY 361 ACATATCATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTTGCCACA 420  
 DB 442 ACATATCATATCTATGAACTCTTAGAGTCAATCAGCCCTCCATTAAGGGGATTTGCCACA 501  
 QY 421 TTTAAACAATATTTGGAGTGTTCGAAGATCGAAGCGACGAGTGGCAGATTTCTGTGAGC 480  
 DB 502 TTTAAACAATATTTGGAGTGTTCGAAGATCGAAGCGACGAGTGGCAGATTTCTGTGAGC 561  
 QY 481 AACCACTTTAGAGCGTGGGAAACTTAGGATGATATGGGAAATGATGAAAGTTCGCG 540  
 DB 562 AACCACTTTAGAGCGTGGGAAACTTAGGATGATGAAAGTTCGCG 621  
 QY 541 CTTACTGTAGAAGGCTATCAAAGTAGCGGAGTCTTAATGTATATAGCAATACACTAAGA 600  
 DB 622 CTTACTGTAGAAGGCTATCAAAGTAGCGGAGTCTTAATGTATATAGCAATACACTAAGA 681  
 QY 601 ATTAACGGTAACCTCTCTCAACTATTAGTAATGACGAGAGCATAACTTTGGATAAAAC 660  
 DB 682 ATTAACGGTAACCTCTCTCAACTATTAGTAATGACGAGAGCATAACTTTAGATAAAAC 741  
 QY 661 AAT 663  
 DB 742 AAT 744

## RESULT 5

AAV30255  
 ID AAV30255 standard; DNA; 871 BP.

XX AC AAV30255;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 18-AUG-1998 (first entry)  
 XX  
 DE DNA encoding a *Bacillus* agaradherens xylanolytic enzyme.  
 XX  
 KW Xylanolytic enzyme; *Bacillus* agaradherens NCIMB 40482; breakdown;  
 KW agricultural waste; alcohol fuel; enzymatic treatment; animal feed;  
 KW release; free pentose sugar; dissolving pulp; cellulose; bio-bleaching;  
 KW wood pulp; lignocellulosic material; animal feed additive; ss.  
 XX  
 OS *Bacillus* agaradherens.

XX FH Key Location/Qualifiers  
 FT CDS 82. 747  
 FT FT /tag= a

XX US5770424-A.

XX PD 23-JUN-1998.

XX PF 16-AUG-1996; 96US-00698978.

XX PR 02-JUL-1993; 93WO-DK000218.

XX PR 30-NOV-1994; 94US-00343600.

XX PR 06-JUN-1995; 95US-00470398.

XX PA (NOVO ) NOVO-NORDISK AS.

XX Outtrup H, Bisgard-Frantzen H, Schuelein M, Olsen AA;

XX PI Jorgensen PL, Dambmann C;

XX DR WPI; 1998-376805/32.

XX DR P-PSDB; AAW60562.



XX DNA construct encoding *Bacillus agaradhaerens* xylanolytic enzyme - and  
PT vectors and *Bacillus* cells containing these, useful for recombinant  
PT production of the enzyme, for use in agricultural waste breakdown and  
PT lignocellulosic material treatment.

XX  
XX  
XX Claim 1; Col 11-14; 10pp; English.

XX  
XX  
XX The present sequence encodes a xylanolytic enzyme of *Bacillus*  
CC *agaradhaerens* NCIMB 40482. Xylanolytic enzymes are used for enzymatic  
CC breakdown of agricultural wastes for production of alcohol fuels,  
CC enzymatic treatment of animal feeds to release free pentose sugars,  
CC manufacturing of dissolving pulps yielding cellulose and bio-bleaching of  
CC wood pulp. they are also used for treatment of lignocellulosic material  
CC e.g. paper and pulp, or as an animal feed additive. (Updated on 25-MAR-  
CC 2003 to correct PI field.) (Updated on 25-MAR-2003 to correct PI field.)  
CC (Updated on 27-AUG-2003 to correct OS field.) (Updated on 17-OCT-2003 to  
CC standardise OS field)

XX  
XX  
XX Sequence 871 BP; 283 A; 171 C; 179 G; 238 T; 0 U; 0 Other;

SO

Query Match	95.7%	Score 634.2	DB 2	Length 871
Best Local Similarity	97.3%	Pred. 2.4e-179		
Matches 645	Conservative 0	Mismatches 18	Indels 0	Gaps 0
Qy	1	CAAAATCGTCACCGCAAAATTCATTTGGCAACACCAAGTGGCTATGATTTATGAATTTTGGAAA	60	
Db				
Qy	82	CAAAATCGTCACCGCAAAATTCATTTGGCAACACCAAGTGGCTATGATTTATGAATTTTGGAAA	141	
Db				
Qy	61	GATAGCGGTGGCTCTGGGACAAATGATTTCTCAATCATGGCGGTACGTTTCAGTGCCCAATGG	120	
Db				
Qy	142	GATAGCGGTGGCTCTGGGACAAATGATTTCTCAATCATGGCGGTACGTTTCAGTGCCCAATGG	201	
Db				
Qy	121	AACAAATGTTAAACAATATTTATTCGGTAAAGGTAAAAATTTCAATGAAGAACACCAACACAC	180	
Db				
Qy	202	AACAAATGTTAAACAATATTTATTCGGTAAAGGTAAAAATTTCAATGAAGAACACCAACACAC	261	
Db				
Qy	181	CAACAAGTGTGTACATGTCATATAAATCTACGAGGCAACTTCCAAACCAAAATGGTAAATGCG	240	
Db				
Qy	262	CAACAAGTGTGTACATGTCATATAAATCTACGAGGCAACTTCCAAACCAAAATGGTAAATGCG	321	
Db				
Qy	241	TATTTATGCGTCTATGGTTGGACTGTTTGACCCCTCTTGTGCAATATTTATTTGTGCAAGT	300	
Db				
Qy	322	TATTTATGCGTCTATGGTTGGACTGTTTGACCCCTCTTGTGCAATATTTATTTGTGCAAGT	381	
Db				
Qy	301	TGGGGCAACTCGCGTCCACAGGAGCAACCGCTTAAGGGGACCATCACTGTTGATGGAGGA	360	
Db				
Qy	382	TGGGGCAACTCGCGTCCACAGGAGCAACCGCTTAAGGGGACCATCACTGTTGATGGAGGA	441	
Db				
Qy	361	ACATATGATATCTACGAGACTCTTTAGAGTCAATCAACCCCTCCATTAAAGGGGATTCGCCACA	420	
Db				
Qy	442	ACATATGATATCTATGAAGTCTTTAGAGTCAATCAACCCCTCCATTAAAGGGGATTCGCCACA	501	
Db				
Qy	421	TTTAAACAATATTGGAGTGTTCGAAGATCGAAACCGCACGAGTGGCACGATTTCTGTCAAGC	480	
Db				
Qy	502	TTTAAACAATATTGGAGTGTTCGAAGATCGAAACCGCACGAGTGGCACGATTTCTGTCAAGC	561	
Db				
Qy	481	AACCACTTTTAGAGCGTGGGAAAACCTTAGGATGAAATATGGGGAAAATGTATGAAGTCCGG	540	
Db				
Qy	562	AACCACTTTTAGAGCGTGGGAAAACCTTAGGATGAAATATGGGGAAAATGTATGAAGTCCGG	621	
Db				
Qy	541	CTTACTGTAGAGGCTATCAAGTAGCGGAAGTGTCAATGTATATAGCAATACACTAAGA	600	
Db				
Qy	622	CTTACTGTAGAGGCTATCAAGTAGCGGAAGTGTCAATGTATATAGCAATACACTAAGA	681	
Db				
Qy	601	ATTTAACGGTAAACCCCTCTCTCAACTATTAGTAAATGACGAGACATACTTTGGATAAAAAC	660	
Db				
Qy	682	ATTTAACGGTAAACCCCTCTCTCAACTATTAGTAAATGACGAGACATACTTTGGATAAAAAC	741	
Db				
Qy	661	AAT 663		
Db				
Qy	742	AAT 744		
Db				

RESULT 6	
AAQ92878	
ID	AAQ92878 standard; DNA; 744 BP.
XX AC	AAQ92878;
XX DT	16-OCT-2003 (revised)
DT	12-FEB-1996 (first entry)
XX DE	Thermostable alkaline endo-1.4-beta-D-xylanase gene.
XX KW	thermostable alkaline endo-1.4-beta-D-xylanase gene; cloning;
KW	polymerase chain reaction; Escherichia coli; EC-3.2.1.8; paper; pulp;
KW	bleaching; db.
XX OS	Bacillus sp; 1-43-3 (CBS 672.93).
XX FN	WO9518219-A1.
XX PD	06-JUL-1995.
XX PF	23-DEC-1994; 94WO-EP004312.
XX PR	24-DEC-1993; 93EP-00203694.
XX PA	(KONN ) GIST-BROCADES NV.
PI	van Solingen P, Williams DP, Iverson S, Farrell RL, Herbes WT;
PI	Van Der Kleij WA, Herweijer MA, Van Beckhoven RPWC, Quax WJ;
PI	Goedegebuur F, Jones BE;
DR	WPI; 1995-246385/32.
DR	P-PSTDB; AAR76551.
XX	Novel xylanase enzyme active at high pH - useful in paper and pulp prodn.
PT	processes.
XX	Claim 4; Page 42-43; 54pp; English.
XX	The sequence encodes an alkaline endo-1.4-beta-D-xylanase (G-type) from
CC	Bacillus sp. 1-43-3. Fragments of the sequence (e.g. the internal
CC	fragments given in AAQ92876 and AAQ92877) may be amplified by polymerase
CC	chain reaction, e.g. using primers with sequences AAQ92866, AAQ92867,
CC	AAQ92868 and AAQ92869. The DNA may be cloned in Escherichia coli using a
CC	plasmid vector for recombinant xylanase production. The xylanase may be
CC	used in the paper and pulp industries, where it produces an increase in
CC	ISO brightness of softwood pulp of at least 1.0 over non-enzymatically
CC	treated pulp in an ECF pulp bleaching process, at pH 9.0 and 65 deg C.
CC	The enzyme may be used in production of paper, board and fluff pulp, and
CC	has low cellulase activity. The increased brightness produced using the
CC	xylanase allows reduction in the amount of bleaching chemicals used.
CC	(Updated on 16-Oct-2003 to standardise OS field)
XX	
SQ	Sequence 744 BP; 248 A; 145 C; 159 G; 192 T; 0 U; 0 Other;
Query Match            94.2%; Score 624.6; DB 2; Length 744;	
Best Local Similarity    96.4%; Pred No. 1.7e-176;	
Matches 639; Conservative    0; Mismatches 24; Indels    0; Gaps    0	
OY	1 CAATTCGTACCGACAATTCCATTGGCAACCAGTGCATGATTTATGAATTTTGAAA 60
DB	79 CAATTCGTACCGACAATTCCATTGGCACC CGCGTGTATGATTATGAATTTTGAAA 138
OY	61 GATACCGTGGCTCTGGGACAATGATTTCTCAATCATGCGGTACGTTCAAGTCCCAATGG 120
DB	139 GATACCGTGGCTCTGGGACAATGATTTCTCAATCATGCGGTACGTTCAAGTCCCAATGG 198
OY	121 AACATGTTAACACATATTTTCGTAAGGTAAAAAATTCATGAAACACAAACACAC 180
DB	199 ATAATGTTAACATATATTTTCGTAAGGTAAAAAATTCATGAAACACAAACACAC 258
OY	181 CAACAAGTTGGTAACATGTCCATAAATCACGGAGCCAATTCCTCAACCAATGGTATGCG 240

```
Db 259 CAACAGTGTGTAACATGCTCCATTAACATATGCGCAACATTCAGCCCAACGGTAATGCG 318
Qy 241 TATTATCCGCTCTATGTTGGACTGTGACCCCTCTTGTGCAATATATATATGTCGACAGT 300
Db 319 TATTATCCGCTCTATGTTGGACTGTGACCCCTCTTGTGCAATATATATATGTCGACAGT 378
Qy 301 TGGGCAACTGGCTGCCACAGGAGCAACGCTTAAGGGACCATCCTGTTGATGGAGGA 360
Db 379 TGGGCAACTGGCTGCCACAGGAGCAACGCTTAAGGGACCATCCTGTTGATGGAGGA 438
Qy 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAGGGGATTTGCCACA 420
Db 439 ACATATGATATCTAGAACTCTTAGAGTCAATCAGCCCTCCATTAGGGGATTTGCCACA 498
Qy 421 TTTAAACAATATGAGTGTTCGAAGATCGAAACGCGACAGTGGCAGCAATTTCTGTGAGC 480
Db 499 TTTAAACAATATGAGTGTTCGAAGATCGAAACGCGACAGTGGCAGCAATTTCTGTGAGC 558
Qy 481 AACCACTTTAGAGCTGGGAAACCTTAGGGATGAATATGGGAAATGTATGAAGTCGCG 540
Db 559 AACCACTTTAGAGCTGGGAAACCTTAGGGATGAATATGGGAAATGTATGAAGTCGCG 618
Qy 541 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA 600
Db 619 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA 678
Qy 601 ATTACGGTAACCCCTCTCTCAACTATTAGTATAGCAGAGACATTAACCTTTGGATAAACA 660
Db 679 ATTACGGTAACCCCTCTCTCAACTATTAGTATAGCAGAGACATTAACCTTTGGATAAACA 738
Qy 661 AAT 663
Db 739 AAT 741
```

## RESULT 7

```
AAV13074
ID AAV13074 standard; DNA; 747 BP.
XX
AC AAV13074;
XX
DT 19-MAY-1998 (first entry)
XX
DE Xylanase activity positive clone DNA SEQ ID NO:11.
XX
KW Bacillus sp; xylanase; glycosyl hydrolase family 11; isolation;
KW microorganism; identification; hybrid DNA; ss.
XX
OS Synthetic.
XX
OS Bacillus sp.
XX
Key Location/Qualifiers
FH 1..747
FT /tag= a
FT /product= "positive clone for xylanase activity"
XX
PN WO9743409-A2.
XX
XX 20-NOV-1997.
XX
XX 12-MAY-1997; 97WO-DK000216.
XX
XX 10-MAY-1996; 96DK-00000562.
XX
XX (NOVO ) NOVO-NORDISK AS.
XX
XX Dalbøge H, Diderichsen B, Sandal T, Kauppinen S;
XX WPI; 1998-008878/01.
XX
XX Isolating novel DNA sequences from microorganisms - without the need for
XX culturing the microorganism.
XX
```

## RESULT 8

```
XX Example 1; Page 35; 72pp; English.
PS
XX The present sequence represents a positive clone for xylanase activity
CC from an example of the present invention. The present invention describes
CC a novel method for providing a novel DNA sequence encoding a polypeptide
CC from a microorganism with an activity of interest. The method comprises:
CC (i) PCR amplification of the DNA with PCR primers with homology to (a)
CC known gene(s) encoding a polypeptide with an activity of interest; (ii)
CC linking the obtained PCR product of a 5' structural gene sequence and a
CC 3' structural gene sequence; (iii) expressing the resulting hybrid DNA
CC sequence; (iv) screening for hybrid DNA sequences encoding a polypeptide
CC with the activity of interest or a related activity; and (v) isolating
CC the hybrid DNA sequence identified in step (iv). This method provides for
CC identification and isolation of sequences from microorganisms without
CC having to cultivate and isolate the microorganism
XX
SQ Sequence 747 BP; 236 A; 168 C; 154 G; 189 T; 0 U; 0 Other;
Query Match 76.6%; Score 507.8; DB 2; Length 747;
Best Local Similarity 85.4%; Pred. No. 1.6e-141;
Matches 566; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
Qy 1 CAATCGTCCCGGACAATTCATTTGGCAACGATGGCTATGATTATGATTTTGGAAA 60
Db 82 CAATCGTCCCGGACAATTCATTTGGCAACGATGGCTATGATTATGATTTTGGAAA 141
Qy 61 GATAGCGTGGCTCTGGGACAATGATTCTCAATCATGCGGTACGTTTCAGTGCCTAATGG 120
Db 142 GATAGCGTGGCTCTGGGACAATGATTCTCAATCATGCGGTACGTTTCAGTGCCTAATGG 201
Qy 121 AACAAATGTTAAACAACATATTATTCCGTAAGAGTAAAAAATTCATATGAAAAACAACAC 180
Db 202 AACAAATGTTAAACAACATATTATTCCGTAAAGTAAAAAATTCATATGAAAAACAACAC 261
Qy 181 CAACAAGTTGGTAACATGTCCTAATAACTACGAGCAACTTCCACCAATGTTAATGCG 240
Db 262 CAACAAGTTGGTAACATGTCCTAATAACTATGGCGCAAACTTCCAGCCAAACGGAATGCG 321
Qy 241 TATTATTCGCTCTATGTTGGACTGTGACCCCTCTTGTCCGAATATATTATGTCGACAGT 300
Db 322 TATTATTCGCTCTATGTTGGACTGTGACCCCTCTTGTCCGAATATATTATGTCGATAGT 381
Qy 301 TGGGGCAACTGGCTGCCACAGGAGCAACGCTTAAGGGACCATCCTGTTGATGGAGGA 360
Db 382 TGGGGCAACTGGCTGCCACAGGAGCAACGCTTAAGGGACCATCCTGTTGACGGGGG 441
Qy 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAGGGGATTTGCCACA 420
Db 442 ACGTATGATATCTACAAGCACCAACAGGTCAATCAGCCATCTATTTCAGGGCACCGCCACC 501
Qy 421 TTTAAACAATATTCGAGTGTTCGAAGATCGAAACGCGACAGTGGCAGCAATTTCTGTACG 480
Db 502 TTTCAATCAGTACTGTCGATTTCGACAGAGCAAGCGGACAGCGGCACTGTCTACAGGCA 561
Qy 481 AACCACTTTAGAGCTGGGAAAACTTAGGAGTAAATATGCGGAAATGTATGAAGTCGCG 540
Db 562 AACCACTTTAATGCTGGGCTGCTCTTGGCATGAATATGGGTGCAATTAACAGATC 621
Qy 541 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA 600
Db 622 CTTGTTACTGAGGGCTACCAATCTACCGGAAGTGTCTAATGTATATAGCAATACACTAAGA 681
Qy 601 ATTACGGTAACCCCTCTCTCAACTATTAGTAAATGACGAGACATAAATTTTGGATAAACA 660
Db 682 ATTACGGTAACCCCTCTCTCAACTATTAGTAAATGACGAGACATAAATTTTGGATAAACA 741
Qy 661 AAT 663
Db 742 AAT 744
```

ADJ34947  
ID ADJ34947 standard; DNA; 1068 BP.  
XX AC ADJ34947;  
XX DT 22-APR-2004 (first entry)  
XX DE DNA encoding xylanase from an environmental sample seq id 163.  
XX DE XX  
XX KW antibacterial; fungicide; thermostable xylanase activity;  
XX KW dough conditioning; beverage production; nutritional supplement;  
KW animal feed; lignin reduction; wood product; xylan; bacterial infection;  
KW fungal infection; coccidiosis; gene; ds.  
XX OS Unidentified.  
OS WO2003106654-A2.  
PN XX  
XX XX  
XX PD 24-DEC-2003.  
XX PF 16-JUN-2003; 2003WO-US019153.  
XX PR 14-JUN-2002; 2002US-0389299P.  
XX PS (DIVE-) DIVERSA CORP.  
XX PA Steer B, Callen W, Healey S, Hazlewood G, Wu D, Blum D;  
PI Esteghlalian A;  
XX XX  
DR WPI; 2004-099016/10.  
DR P-PSDB; ADJ34948.  
XX XX  
XX PT Novel xylanase recombinant polypeptide useful for improving textile  
PT texture, treating paper, eliminating microorganisms.  
XX PS Claim 1; SEQ ID NO 163; 570pp; English.  
XX XX  
XX The invention describes an isolated or recombinant polypeptide (I),  
CC having 50% or more identity to 190 300-1200 residue amino acid sequences  
CC (S1), given in the specification, over a region of 100 or more residues  
CC and the polypeptide as thermostable xylanase activity. (I) is useful for:  
CC dough conditioning; beverage production; as a nutritional supplement in  
CC animal feed; reducing lignin in a wood or a wood product; and for  
CC eliminating and protecting animals from a microorganism comprising xylan.  
CC The polynucleotide (II) encoding (I) is useful for amplifying nucleic  
CC acid encoding a polypeptide having a xylanase activity which involves  
CC amplification of a template nucleic acid with a primer pair capable of  
CC amplifying (II) or its subsequence. (I) is useful for treating and  
CC preventing bacterial infection and fungal infection e.g. coccidiosis.  
CC This sequence encodes xylanase protein isolated from an environmental  
CC sample.  
XX XX  
SQ Sequence 1068 BP; 303 A; 249 C; 279 G; 237 T; 0 U; 0 Other;  
  
Query Match 56.0%; Score 371.2; DB 12; Length 1068;  
Best Local Similarity 76.8%; Pred.No. 1.5e-100;  
Matches 467; Conservative 0; Mismatches 138; Indels 3; Gaps 1;  
  
QY 1 CAATCGTCACCGAGCAATTCCATTTGGCACCACGATGGCTATGATTGAATTTTGGAAA 60  
Db 79 CAGACGNATCACCGAAATTCGGTCGGTGACCATGACGGTTATGACTATGAACTACTGGAAG 138  
  
QY 61 GATAGCGGTGGCTCTGGGACAATGATCTCAATCATCTGCGGGTAGCTTCAGTCCCATAATGG 120  
Db 139 GACAGCGGGGAA---TGCAACTATGTTCTCTCGGTAGTGGCGGTACGTTTCAGTCCGAGTGG 195  
  
QY 121 AACAAATGTTAACACATATTTATTCCTGTAAGGTAAAAAATTCATGAACACACACACAC 180  
Db 196 AGCAATATCANATAATTTCTGTTCCGTAAGGCAGAGTTCAATGAGACGAGACCCAT 255  
  
QY 181 CAACAAGTTGGTAAACATGTCATAAATACAGGAGGCCAATTCACACCAAATGGTAATGG 240  
256 CAGCAATTTGGAACATTTTCCAATCTATGGTGGCCACTACCAACCGAATGGCAATTCG 315

CC animal feed, reducing lignin in a wood or a wood product; and for  
CC eliminating and protecting animals from a microorganism comprising xylan.  
CC The polynucleotide (II) encoding (I) is useful for amplifying nucleic  
CC acid encoding a polypeptide having a xylanase activity which involves  
CC amplification of a template nucleic acid with a primer pair capable of  
CC amplifying (II) or its subsequence. (I) is useful for treating and  
CC preventing bacterial infection and fungal infection e.g. coccidiosis.  
CC This sequence encodes xylanase protein isolated from an environmental  
CC sample.

Db 830 TGGGCGAGTGGCGTCCACCTGGAGCAACACCCAAAGGAACCATCACACAGTGGATGCA 889  
 Qy 358 GGAACATATGATATACAGAGCTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATGGC 417  
 Db 890 GGTACTTATGAATATATGAACCTACCCGGGTAAATCAGCCTTCCATCGATGGAACCTGG 949  
 Qy 418 ACATTTAAACATATTTGGAGTGTTCGAAGATCGAAACGACAGTGGCAGGATTTCTGTC 477  
 Db 950 ACATTTCAACATATTTGGAGTGTTCGATATCAAGAGAACCAAGCGGACATATCTGTC 1009  
 Qy 478 AGCAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATCGGGAAATGTATGAAGTC 537  
 Db 1010 ACTGACATTTTAAACAGTGGGAAAGAAATGGGCATCGGAATGGTAAAGATGTATGAAGTT 1069  
 Qy 538 GGCCTTACTGTAGAGGCTATCAAGTACGGAAGTGTCTAATGTATATAGCAATACACTA 597  
 Db 1070 GCTCTTACCGTTGAAGTGTATCAGAGCAGTGGGTACGCTTAATGTATACAAGAAATGAAATC 1129  
 Qy 598 AGAATTAACGGTAACCC 614  
 Db 1130 AGAATAGGTGCAATCC 1146

## RESULT 11

AAQ80923 standard; DNA; 1022 BP.

XX ID AAQ80923 standard; DNA; 1022 BP.  
 AC AAQ80923;  
 XX 02-AUG-1995 (first entry)  
 DT  
 XX B. pumilus xylanase gene.  
 DE  
 XX Xylanase; endo-1,4-beta-D-xylanase; wood pulp; pulping; biobleaching;  
 KW bleaching; Bacillus licheniformis; ss.  
 XX  
 OS Bacillus pumilus.

Key Location/Qualifiers  
 5'UTR 1..185  
 /tag= a  
 /note= "claim 12"  
 misc\_difference 107  
 /tag= b  
 /note= "base n at position 107 is not identified in the specification"  
 CDS 186..869  
 /tag= c  
 sig\_peptide 186..266  
 /tag= d  
 /note= "claim 11"  
 mat\_peptide 267..866  
 /tag= e  
 /EC number= "3.2.1.8"  
 /note= "claim 10"

GB2279955-A.

18-JAN-1995.

15-JUL-1993; 93GB-00014780.

15-JUL-1993; 93GB-00014780.

(SOLV ) SOLVAY & CIE.

Amory A, Andre C, De Buyl E, Detroz R, Lahaye A, Ledoux P;

WPI; 1995-039214/06.

P-PSDB; AAR68849.

Purified xylanase from Bacillus pumilus PRL B12 - esp. produced in

PT transformed Bacillus licheniformis, and related DNA, vectors, etc., used  
 PT for pre-treatment of wood pulp to reduce chlorine or ozone consumption in  
 PT subsequent bleaching.

XX Claim 13; Fig 1a-1b; 97pp; English.

CC A Bacillus pumilus PTL B12 (ATCC 55443) gene library was screened for  
 CC recombinant plasmids carrying the xylanase gene. A chromosomal fragment  
 CC obtained from isolate pBPXI was subcloned and expressed in Escherichia  
 CC coli JM109. The sequence of the xylanase gene (as a 1022 bp Sau3AI  
 CC fragment) carried by a selected transformant is given in AAQ80923

XX Sequence 1022 BP; 329 A; 197 C; 229 G; 266 T; 0 U; 1 Other;

SQ Query Match 38.5%; Score 255.2; DB 2; Length 1022;  
 Best Local Similarity 65.5%; Pred. No. 8.4e-66;  
 Matches 406; Conservative 0; Mismatches 208; Indels 6; Gaps 2;

Qy 13 GACAAATCCATTTGGCAACGAGTGGCTATGATATGAATTTTGGAAAGATAGCGGTGGC 72  
 Db 279 GATAATAGGATAGGACACACAGCGGATACGATTTTGAATTTATGGAAGGATTAC---GGA 335  
 Qy 73 TCTGGCAATGATCTCAATCATGGCGGTAGCTTCAGTGGCCCAATGGAACAATGTTAAC 132  
 Db 336 AATACCTCGATGACACTCAATAACGGCGGGGCAATTTAGTGAAGCTGGAACAATATGGA 395  
 Qy 133 AACATATTATTCCGTAAGAGTAAATAATTCATATGAACACAAACACACCAACAGTTGTT 192  
 Db 396 AATGCTTTATTTTCGAAAGAGAAAGAGTTTGNATCCACTAAACATCATCATCACTTGC 455  
 Qy 193 AACATGTCTCAATAAATACGGAGCCAACTTCCAAACCAATGGTAATGCGTATTTATGCGTC 252  
 Db 456 AACATCTCCATCAACTACAAACGAGCCTTTAAACCGGCGGGAATTCCTATTTATGTC 515  
 Qy 253 ATGTTGGAGTGTGGACCTCTTTCGAAATATATATGTCGACAGTTTGGGCGCACTCG 312  
 Db 516 TATGGCTGGACACAATCTCCATTAGCTGAATACTACATTTGTTGAGTCATGGGCGCATAT 575  
 Qy 313 CGTCCACAGGAGCAACGCTTAAGGGGACCATCACTGTTGTATGGAGGAACATATGATATC 372  
 Db 576 CGTCCACAGG---AACGTATTAAGGATCATTTTATGCCGATGGAGGCACATATGACATA 632  
 Qy 373 TAGGAGACTTTTAGAGTCAATCAACCTCCATTAAGGGGATTTGCCACATTTAAACAATAT 432  
 Db 633 TATGAAACGCTCCGTGTCAATCAGCCTTCTATCATTTGGAGAGCGCTACCTTCAACAATAT 692  
 Qy 433 TGGAGTGTGGAAGATCGAAACGACGAGTGGCAGCATTTCTGTGACGAACCACTTTTGA 492  
 Db 693 TGGAGTGTACGTCAAAACAAACGCAACGGAACCGGTCTCCGTGAGTGGAGCATTTTAA 752  
 Qy 493 GCGTGGGAAACTTTAGGATGAATATGGGAAATATGTAAGTCCGCTTACTGTAGAA 552  
 Db 753 AAATGGGAAAGCTTAGGCAATGGCAATGGGAAATATGTAAGAACAGCATTTAACTGTAGAA 812  
 Qy 553 GGCTATCAAAGTAGCGGAAGTCTAATGTATATAGCAATACACTAAGAAATTAACGGTAAC 612  
 Db 813 GGCTACCGAAGCAACGGAAGTGGCAATGTGATGACGAATCAGCTGATGATTCGATAAAG 872  
 Qy 613 CCTCTCTCAACTATTAGTAA 632  
 Db 873 CATATGAAAAAGCCAGCAA 892

## RESULT 12

ADJ34969  
 ID ADJ34969 standard; DNA; 684 BP.

XX ADJ34969;

XX AC  
 XX DT 22-APR-2004 (first entry)

XX DNA encoding xylanase from an environmental sample seq id 185.

XX

KW antibacterial; fungicide; thermostable xylanase activity;  
 KW dough conditioning; beverage production; nutritional supplement;  
 KW animal feed; lignin reduction; wood product; xylan; bacterial infection;  
 KW fungal infection; coccidiosis; gene; ds.  
 XX Unidentified.  
 XX WO2003106654-A2.  
 XX 24-DEC-2003.  
 XX 16-JUN-2003; 2003WO-US019153.  
 XX 14-JUN-2002; 2002US-0389299P.  
 XX (DIVE-) DIVERSA CORP.  
 XX Steer B, Callen W, Healey S, Hazlewood G, Wu D, Blum D;  
 PI Esteghlalian A;  
 XX WPI; 2004-099016/10.  
 DR P-PSDB; ADJ34970.  
 XX Novel xylanase recombinant polypeptide useful for improving textile  
 PT texture, treating paper, eliminating microorganisms.  
 XX Claim 1; SEQ ID NO 185; 570pp; English.  
 XX The invention describes an isolated or recombinant polypeptide (I),  
 CC having 50% or more identity to 190 300-1200 residue amino acid sequences  
 CC (S1), given in the specification, over a region of 100 or more residues  
 CC and the polypeptide as thermostable xylanase activity. (I) is useful for:  
 CC dough conditioning; beverage production; as a nutritional supplement in  
 CC animal feed; reducing lignin in a wood or a wood product; and for  
 CC eliminating and protecting animals from a microorganism comprising xylan.  
 CC The polynucleotide (II) encoding (I) is useful for amplifying nucleic  
 CC acid encoding a polypeptide having a xylanase activity which involves  
 CC amplification of a template nucleic acid with a primer pair capable of  
 CC amplifying (II) or its subsequence. (I) is useful for treating and  
 CC preventing bacterial infection and fungal infection e.g. coccidiosis.  
 CC This sequence encodes xylanase protein isolated from an environmental  
 CC sample.  
 XX SQ Sequence 684 BP; 221 A; 129 C; 158 G; 176 T; 0 U; 0 Other;  
 SQ Query Match 38.5%; Score 255; DB 12; Length 684;  
 Best Local Similarity 66.8%; Pred. No. 8.2e-66;  
 Matches 395; Conservative 0; Mismatches 190; Indels 6; Gaps 2;  
 QY 13 GACAAATCCATTGGCAACCAACGATGGCTATGATTTATGAATTTTGGAAAGATAGCGTGGC 72  
 DB 94 GATAATGAGATAGGACACACACAGCGGATACGATTTTGAATTAATGAAGATTAC---GGA 150  
 QY 73 TCTGGACAAATGATTCTCAATCATCGCGGTAGCTTTCAGTGCCTCAATGGAACAATGTTAAC 132  
 DB 151 AATACCTCGATGACACTCAATAACGCGGGCATTTAGTGAAGCTGGAACAATATGGA 210  
 QY 133 AACATATTTCGTAAAGCTAAATAATTCATGAACACACACACCAACCAAGTGGT 192  
 DB 211 AATGCCCTTATTTGGAAGGAAGAAGTTTGATTCACATAAAACTCATCATCACTTGGC 270  
 QY 193 AACATGCTCAATAAATACGAGCAACTTCCAAACCAATGTTAATCGGTATTTATGCGTC 252  
 DB 271 AACATCTCCATCACTACACAGCAGCTTTTACCCGGGGGGAATTCCTATTATGTC 330  
 QY 253 TATGTTGACGTGTTGACCCCTCTTGTGCAATATATATTGTGCGACAGTTGGGGCAACTGG 312  
 DB 331 TATGGCTGACACAAATCTCCATTAGCTGAATACATCATTTGTTGAGTTCATGGGGCACATAT 390  
 QY 313 CGTCCACAGGACGACGCTTAAGGGACCATCCTGTTGATGGAGGAACATATGATATC 372  
 DB 391 CGTCCAAACAGG---AACGTATAAAGGATCATTTTATGCCGATGGAGGCACATATGACATA 447

QY 373 TACGAGACTCTTAGAGTCATCAACCTCCATTAAAGGGATTGCCACATTAAACAATAT 432  
 DB 448 TATGAACAGCTCCGTGTCATCAGCTTCTTATCATTTGGAGACGCTTACCTTCAACAATAT 507  
 QY 433 TGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGATTCTTGTGACGAACCACTTTAGA 492  
 DB 508 TGGAGTGTACGTCAACCAAAACGCAACGCGAACTGTTTCCGTGAGTGAAGCAATTTAAA 567  
 QY 493 GCGTGGGAAAACCTTAGGATGAATATGGGGAATATGTATGAAGTCCGGCTTACTGTAGAA 552  
 DB 568 AAATGGGAAAAGCTTAGGCATGCCAAATGGGAAAATGTATGAACAGCATTAATCTGTAGAA 627  
 QY 553 GCGTATCAAAAGTAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGAAAT 603  
 DB 628 GGTACCGAAGCAACGGAAGTGGCAATGTCTATGCAATCAGCTGATGATT 678  
 RESULT 13  
 AAT90972  
 ID AAT90972 standard; DNA; 1190 BP.  
 XX AC AAT90972;  
 XX DT 22-MAY-1998 (first entry)  
 XX DE Nucleotide sequence encoding an enzyme with xylanase activity.  
 XX KW xynB gene; xylanase; enzyme; Dictyoglomus thermophilum strain Rt46B.1;  
 KW G-Xylanase; beta -1,4-xylanase activity; activity; thermal stability;  
 XX biological bleaching; cellulose product; paper pulp; ss.  
 XX OS Dictyoglomus thermophilum.  
 XX FH Key Location/Qualifiers  
 FT CDS 67..1149  
 FT /\*tag= a  
 FT /\*product= "xylanase"  
 FT sig\_peptide 67..138  
 FT /\*tag= b  
 FT mat\_peptide 139..1146  
 FT /\*tag= c  
 XX WO9736995-A2.  
 XX 09-OCT-1997.  
 XX 27-MAR-1997; 97WO-NZ0000042.  
 XX 29-MAR-1996; 96NZ-00286296.  
 XX (PACI-) PACIFIC ENZYMES LTD.  
 XX Bergquist PL, Gibbs MD, Morris D;  
 XX WPI; 1997-503090/46.  
 DR P-PSDB; AAW30267.  
 XX Dictyoglomus thermophilum xylanase gene - used for producing thermally  
 PT stable enzymes for the bleaching of cellulose products, especially paper  
 PT pulp.  
 XX Disclosure; Fig 3; 34pp; English.  
 XX The present sequence represents a Rt46B.1 xynB gene and encodes a xylanase  
 CC enzyme. The present sequence is derived from a gene contained within  
 CC Dictyoglomus thermophilum strain Rt46B.1. The xylanase is contained  
 CC within the family of enzymes known as G-Xylanases, and has beta -1,4-  
 CC xylanase activity. The enzyme has high activity and high thermal  
 CC stability with optimum activity at 85 degrees Celsius and pH 6.5. The  
 CC xylanase enzyme is used for the biological bleaching of cellulose  
 CC products, especially paper pulp. Use of the enzyme ensures that waste  
 CC streams from the biological bleaching will include less toxic material  
 CC XX

SQ Sequence 1190 BP; 386 A; 197 C; 277 G; 330 T; 0 U; 0 Other;  
Query Match 32.0%; Score 212.2; DB 2; Length 1190;  
Best Local Similarity 62.4%; Pred. No. 6.8e-53;  
Matches 369; Conservative 0; Mismatches 213; Indels 9; Gaps 2;  
QY 3 AATCGTCACCGCAATTCCTCAATGCGGTCAGTTCAGTCCCAATGGNA 122  
DB 150 AACACTAACAGTAATGCAAGCGGTACTTTGATGGCTACTATGAACATGGAAGA 209  
QY 63 TAGCGGTGGCTCGGACAAATGATTCATCATGCGGTCAGTTCAGTCCCAATGGNA 122  
DB 210 TA---CAGGAATACAACTGACTGTATACACAAAGGAAGTTTGTAGTGTGAGTGGAG 266  
QY 123 CAATGTTAACACATATTTTCGGTAAAGGTAAATTCATGAAACACAAACACCA 182  
DB 267 CAATATAAACAATGCAATATTTCAGAACAGGTAAAGTACAAC-----CAAAACTGGCA 320  
QY 183 ACAAGTTGGTAAATGTCATTAACACTACGAGGCAACTTCCAAACCAATGGTAATGCGTA 242  
DB 321 GTCAATAGGCACTATTAGAACTACCTACTCAGGCCACATATATCTTAATGGTAATCTCTA 380  
QY 243 TTTATCGCTCTATGTTGCACTGTTGACCTCTTGTGCAATATATATTTGTGACAGTTG 302  
DB 381 CTTATGCTATGTTGGTGGTCTACTAATCTTTAGTAGAGTTTATACATTTAGAAAGTTG 440  
QY 303 GGGCAACTGGCTCCACAGGAGCAAGCGCTAAGGGGACATCAGTCTGTTGATGGAGAAC 362  
DB 441 GGGTAATTTGGCTCCACAGGTCACCTCTCTTGGACAGGTTACTATCGACGGTGGTAC 500  
QY 363 ATATGATATCTACGAGCTCTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCCATTT 422  
DB 501 CTATGACATTTACAGAACTACCGTGTAAATCAGCCATCTATTGTGCGTACAGTACTTT 560  
QY 423 TAAACAATATTGAGTGTTCGAAGATCGAAACGACAGTGGCAAGATTTCTGTACAGAA 482  
DB 561 TGATCAATATTGAGTGTGAAGACATCTAAGAGAACAGTGGAAACAGTCACTGTAAACAGA 620  
QY 483 CGACTTTAGAGCTGGGAAACCTTAGGATGAATATGGGGAATATGTAAGTCCGCT 542  
DB 621 TCACCTTTAGGCAATGGGCAATAGAGGTTTAAACCTTGGTACTATTGATCAAAATTA 680  
QY 543 TACTGTAGAGGCTATCAAGAGTAGCGAGTGTCTAATGTATATAGCAATAC 593  
DB 681 TTGTGTTGAAGGATATCAAGACAGTGTGTCGGCTAATATACACAAATATAC 731

RESULT 14

ID AAT08142 standard; DNA; 1244 BP.  
XX  
AC AAT08142;  
XX  
DT 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 12-MAY-1996 (first entry)

Xylanase xynD gene.

XX Xylanase D; xynD gene; thermostable enzyme; paper; pulp; lignin;  
KW delignification; xylan; genan; bleaching; ds.  
XX Bacteria.

OS  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1107  
FT /\*tag= a

XX WO9534662-A1.

XX 21-DEC-1995.

XX 14-JUN-1995; 95WO-EP002299.

XX 14-JUN-1994; 94EP-00201699.  
PR (KONN ) GIST-BROCADES BV.  
XX Groenberg V, Williams DP, Iverson S, Forster S, Moody D;  
PI Farrel RL, Bergquist PL, Daniel RM, Morgan HW, Quax WJ;  
PI Herweijer MA, Jones BE;  
XX WPI; 1996-049690/05.  
DR P-PSDB; AAR87012.  
XX Novel xylanase(s) having activity at 80 deg.C. or higher - obtained from  
PT anaerobic thermophilic bacteria and used in paper and pulp production  
PT processes.  
XX Example 9; Page 56-58; 77pp; English.  
XX A full-length xynD gene (AAR08142) codes for a thermostable G-type  
CC xylanase (AAR07012) useful in the pulp and paper industries. The gene was  
CC obtd. by genomic walking PCR of DNA from an extremely thermophilic  
CC bacterium, strain TGA56 (CBS 213.94), isolated from a New Zealand hot  
CC spring. The gene can be inserted into a vector and used for the prodn. of  
CC recombinant xylanase D in microbial host cells, esp. Escherichia coli.  
CC (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to  
CC correct OS field.)  
XX Sequence 1244 BP; 380 A; 215 C; 279 G; 360 T; 0 U; 10 Other;  
SQ  
Query Match 30.9%; Score 204.8; DB 2; Length 1244;  
Best Local Similarity 61.0%; Pred. No. 1.1e-50;  
Matches 367; Conservative 2; Mismatches 224; Indels 9; Gaps 2;  
QY 8 TCACCGCAATTCATTCGCAACCAAGATGGTATGATTTATTAATTTGGAAAGTAGCG 67  
DB 71 TTACCTCTAATGCAACTGGGACATACGATGGTTACTACTACGAGTTGTGGAAGGACACAG 130  
QY 68 GTGGCTCTGGGCAATGATTTCTCAATCATGGCGGTACGTTCACTGCGCAATGGAACAATG 127  
DB 131 GGAATACT---ACCATGACAGTTGACACAGGAGGAAGATTTAGTCTCAGTGAGTAACA 187  
QY 128 TTAAACAACATATATTTCGGTAAAGGTAAATAATTCATGAAACACAAACACCAACAAG 187  
DB 188 TTAACAATGCACTCTTCAGAACAGGTAAAGTTAGCACTGCATGGAATCAGC----- 241  
QY 188 TTGGTAAACATGTCCTAACTACGGAGCCAACTTCAACCAATGTAATGCGTATTTAT 247  
DB 242 TTGGGACTGTAAAGATTACCTACTCTGCTACCTACAATCCAATGGCAATTCCTATCTCT 301  
QY 248 GCGTCTATGTTGGAGCTGTTGACCTCTTGTGCAATATATATTTGTCGACAGTGGGGCA 307  
DB 302 GCATTTATGATGGTCAAGAAATCACTGTTGAAATTTATATCGTTGAAAGCTGGGCT 361  
QY 308 ACTGGCGTCCACGAGGCAACCGCTTAAGGGACCAATCACTGTTGATGAGGAAACATATG 367  
DB 362 CATGGCTCGCGCCGGGGCAACGTCACCTGGGCACTGTAACAATTTGATGGAGCAACATATG 421  
QY 368 ATATCTACGAGACTCTTAGAGTCAATCAACCTTCAATTAAGGGGATTCGCAATTTAAAC 427  
DB 422 ATATTTATAAGACAACCTCGTGTAAATCAGCCACTCTATCGAAGGAACAAAGAACATTTGATC 481  
QY 428 AATATTGGAGTGTTCGAAGATCGAACGCAAGTGGCAGGATTTCTGTGACGAACCACT 487  
DB 482 AGTACTGGAGTGTTAGGACATCAAGAGAACAGTGGTACTGTTACTGTAACTGATCATTT 541  
QY 488 TTAGAGCGTGGGAAACCTTTAGGATGAATATATGGGGAATATGTAAGTCCGCTTACTG 547  
DB 542 TCAAGCATGGGCTGCAAAAGGTTTGAACCTGGGTACAATTTGACCAAGATTTACACTCTGTG 601  
QY 548 TAGAAGGCTATCAAGTAGCGGAAGTCTAATGTATATAGCATATACACTAAGAAATTAACG 607  
DB 602 TGGAGGTTACCARAGCAGCGGCTCAGCAATATAACACAGAAATATCTTACTATTGGTG 661



QY 608 GT 609  
Db 662 GT 663

RESULT 15  
ADJ35011  
ID ADJ35011 standard; DNA; 747 BP.  
XX AC  
XX ADJ35011;  
XX  
XX 22-APR-2004 (first entry)  
XX  
XX DNA encoding xylanase from an environmental sample seq id 227.  
XX  
XX antibacterial; fungicide; thermostable xylanase activity;  
KW dough conditioning; beverage production; nutritional supplement;  
KW animal feed; lignin reduction; wood product; xylan; bacterial infection;  
KW fungal infection; coccidiosis; gene; ds.  
XX  
XX Unidentified.  
XX  
XX WO2003106654-A2.  
XX PN  
XX 24-DEC-2003.  
XX PD  
XX  
XX PF 16-JUN-2003; 2003WO-US019153.  
XX  
XX PR 14-JUN-2002; 2002US-0389299P.  
XX  
XX (DIVE-) DIVERSA CORP.  
XX PA  
XX Steer B, Callen W, Healey S, Hazlewood G, Wu D, Blum D;  
XX PI Esteghlalian A;  
XX  
XX WPI; 2004-099016/10.  
XX DR P-PSDB; ADJ35012.  
XX  
XX Novel xylanase recombinant polypeptide useful for improving textile  
PT texture, treating paper, eliminating microorganisms.  
XX  
XX Claim 1; SEQ ID NO 227; 570pp; English.  
XX  
XX The invention describes an isolated or recombinant polypeptide (I),  
CC having 50% or more identity to 190 300-1200 residue amino acid sequences  
CC (S1), given in the specification, over a region of 100 or more residues  
CC and the polypeptide as thermostable xylanase activity. (I) is useful for:  
CC dough conditioning; beverage production; as a nutritional supplement in  
CC animal feed; reducing lignin in a wood or a wood product; and for  
CC eliminating and protecting animals from a microorganism comprising xylan.  
CC The polynucleotide (II) encoding (I) is useful for amplifying nucleic  
CC acid encoding a polypeptide having a xylanase activity which involves  
CC amplification of a template nucleic acid with a primer pair capable of  
CC amplifying (II) or its subsequence. (I) is useful for treating and  
CC preventing bacterial infection and fungal infection e.g. coccidiosis.  
CC This sequence encodes xylanase protein isolated from an environmental  
CC sample.

Search completed: November 10, 2004, 17:09:29  
Job time : 424.566 secs

XX SQ Sequence 747 BP; 174 A; 222 C; 244 G; 107 T; 0 U; 0 Other;  
XX  
XX Query Match 26.8%; Score 177.8; DB 12; Length 747;  
XX Best Local Similarity 66.1%; Pred. NO. 1.1e-42;  
XX Matches 257; Conservative 0; Mismatches 132; Indels 0; Gaps 0;  
XX  
QY 193 AACATGTCATAACTACGAGGCAACTTCCACCAATGGTAATCGGTATTTATCGCTC 252  
Db 340 AATCAGACGGTGACCTTACCGGCAAACTACCGCGGACCGCAATTCATCTGTGCGPA 399  
QY 253 TATGTTGGACTCTTGACCCCTTTGTCGAATATTATTTGTCAGAGTTGGGCAACTGG 312  
Db 400 TATGGATGGACGCAAAACCCCTCGTCGAATACATCGTCGATAGTGGGCGAGCTGG 459  
QY 313 CGTCCACGAGGCAACGCCCTAAGGGGACCATCATCTGTTGATGGAGGAACATATGATATC 372



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 16:24:32 ; Search time 351.127 Seconds  
(without alignment)  
8406.125 Million cell updates/sec

Title: US-09-909-207-27  
Perfect score: 81  
Sequence: 1 ATGAATTGAAAGATTGAG.....CGGCTGTCCGGCTCATCGG 81

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30.6	37.8	208	8 AQ08988	GSSTC0102
2	30.6	37.8	348	9 CC867970	NDL.12311
3	30.2	37.3	449	6 CD480598	eca01-23m
4	30.2	37.3	451	4 BJ396832	BJ396832
5	29.8	36.8	663	2 BB628216	BB628216
6	29.8	36.8	1533	3 AK035636	Mus muscu
7	29.6	36.5	684	1 AV609193	AV609193
8	29.6	36.5	748	5 BX083182	EX083182
9	29.6	36.5	804	2 BE052262	GA_Ea003
10	29.6	36.5	1050	9 CNS06ROG	AL412214 T7 end of
11	29.4	36.3	706	8 BH091215	RPCI-24-3
12	29.2	36.0	498	7 CR375533	CR375533
13	29.2	36.0	623	6 CA381308	CA381308
14	29.2	36.0	780	5 BX859701	EX859701
15	29.2	36.0	824	7 CN594115	TTE000128
16	29	35.8	475	8 BZ459067	BONA064TR
17	29	35.8	480	5 BY364919	BY364919
18	29	35.8	485	5 BY365129	BY365129
19	29	35.8	588	2 BF556005	BF556005
20	29	35.8	631	8 BH440455	BH440455
21	29	35.8	762	8 BZ705274	PUBL880TD
22	29	35.8	765	7 CR371060	CR371060
23	29	35.8	810	7 CK597065	AGENCOURT
24	29	35.8	855	8 CC357397	PUHRM53TD

25	28.8	35.6	248	1 AA072982	AA072982
26	28.8	35.6	335	1 AA823781	AA823781
27	28.8	35.6	343	2 BF462060	BF462060
28	28.8	35.6	531	6 CA751489	CA751489
29	28.8	35.6	538	7 CN715895	CN715895
30	28.8	35.6	542	7 CN717635	CN717635
31	28.8	35.6	546	7 CN724716	CN724716
32	28.8	35.6	549	7 CN681450	CN681450
33	28.8	35.6	588	6 CD551944	CD551944
34	28.8	35.6	615	2 BB586000	BB586000
35	28.8	35.6	629	5 BQ770457	BQ770457
36	28.8	35.6	639	7 CF535665	CF535665
37	28.8	35.6	655	7 CF735229	CF735229
38	28.8	35.6	669	5 BM963173	BM963173
39	28.8	35.6	672	6 CB247072	CB247072
40	28.8	35.6	676	5 BU704409	BU704409
41	28.8	35.6	680	7 CN538366	CN538366
42	28.8	35.6	684	7 CK636620	CK636620
43	28.8	35.6	688	5 BQ442114	BQ442114
44	28.8	35.6	700	7 CN533101	CN533101
45	28.8	35.6	713	7 CF535405	CF535405

## ALIGNMENTS

RESULT 1  
AQ08988  
LOCUS GSSTC010294 Trypanosoma cruzi random genomic library Trypanosoma  
DEFINITION cruzi genomic clone GS1N1, genomic survey sequence.  
ACCESSION AQ08988  
VERSION AQ08988.3 GI:10130354  
KEYWORDS GSS.  
SOURCE Trypanosoma cruzi  
ORGANISM Trypanosoma cruzi  
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;  
Trypanosoma; Schizotrypanum.  
REFERENCE  
1 (bases 1 to 208)  
Aguero, F., Verdun, R., Frasch, A.C.C. and Sanchez, D.O.  
A random sequencing approach for the analysis of the trypanosoma  
cruzi genome: general structure, large gene and repetitive DNA  
families, and gene discovery  
Genome Res. 10 (12), 1996-2005 (2000)  
20568489  
11116034  
On Sep 14, 2000 this sequence version replaced gi:9377651.  
Contact: Sanchez D.O.  
Instituto de Investigaciones Biotecnologicas (Univ. Nac. de Gral  
San Martin)  
Av. Gral Paz S/N, INTI, Edificio 24, B 1650 KNA, San Martin, Buenos  
Aires, Argentina  
Tel: (54-11) 4580/7255/7  
Fax: (54-11) 4752-9639  
Email: dsanchez@ib.unsam.edu.ar  
Sequences were basecalled with phred and vector was masked with  
crossmatch (see http://genome.washington.edu). Sequences were then  
trimmed from both ends to remove low quality bases and masked  
vector.  
Seq primer: T7  
Class: shotgun.

FEATURES  
source  
Location/Qualifiers  
1..208  
/organism="Trypanosoma cruzi"  
/mol\_type="genomic DNA"  
/strain="CL-Brener"  
/db\_xref="taxon:5693"  
/clone="GS1N1"  
/cell\_type="epimastigote"  
/clone\_lib="Trypanosoma cruzi random genomic library"  
/note="Vector: pBS(-) (Stratagene); T. cruzi DNA was  
randomly sheared using a nebulizer and the 1 to 2 kb range  
was gel purified and cloned into the dephosphorylated

```

ORIGIN
  HincII site of the vector"

Query Match      37.8%; Score 30.6; DB 8; Length 208;
Best Local Similarity 65.2%; Pred. No. 33;
Matches 45; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 11 AAAGATTGAGCTGTGTTTGTGATGTATGATTGCTGCTGACACTGACGCTGTGC 70
    |||||
Db 68 ACATGTGAGGCTGTGTTTGTGATGTGTTCTTTCTGTTGACCCGACGCTGTGTC 127
    |||||

Qy 71 CGGCTCATG 79
    |||||
Db 128 TGGATTGTG 136

RESULT 2
CC867970/c
LOCUS
DEFINITION
  ND1.123116.T7 Notre Dame Liverpool Aedes aegypti genomic clone
  Notre Dame Liverpool-123116, genomic survey sequence.
ACCESSION
  CC867970
VERSION
  CC867970.1 GI:33227980
KEYWORDS
  GSS
ORGANISM
  Aedes aegypti (yellow fever mosquito)
  Aedes aegypti
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes;
  Stegomyia.
REFERENCE
  1 (bases 1 to 348)
  Loftus,B., Shetty,J., Knudson,D. and Severson,D.
  BAC end sequencing of Aedes aegypti
  Unpublished (2003)
JOURNAL
  Contact: Brendan Loftus
  Department of Eukaryotic Genomics
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-3543
  Fax: 301-838-0208
  Email: enta@tigr.org
  Library was provided by David Severson
  Seq primer: 17
  Class: BAC ends.
FEATURES
  source
    1..348
    /organism="Aedes aegypti"
    /mol_type="genomic DNA"
    /strain="Liverpool"
    /db_xref="taxon:7159"
    /clone="Notre Dame Liverpool-123116"
    /clone_lib="Notre Dame Liverpool"
    /note="Vector: pSCBAC1; Site 1: Hind III; The library was
    prepared from whole body tissue of newly hatched L1 larvae
    by David Severson at the University of Notre Dame and
    Hongbin Zhang"

ORIGIN
  Query Match      37.8%; Score 30.6; DB 9; Length 348;
  Best Local Similarity 73.6%; Pred. No. 36;
  Matches 39; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 ATGATTTTGAAGATTGAGGCTGTGTTTGTGATGTGATTGATTTGTGCT 53
    |||||
Db 115 ATGATTTTGAATATTTAGGCTGTGGATGTGATATATATCGCTGTGAT 63
    |||||

RESULT 3
CD480598
LOCUS
DEFINITION
  eca01-23msl-g10 Eca01 Eschscholzia californica cDNA clone
  eca01-23msl-g10 5', mRNA sequence.
ACCESSION
  CD480598
VERSION
  CD480598.1 GI:31401866

ORIGIN
  Query Match      37.3%; Score 30.2; DB 6; Length 449;
  Best Local Similarity 65.7%; Pred. No. 51;
  Matches 44; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 6 TTTGAAAAGATTGAGGCTGTGTTTGTGATGTGATTGTTGCTGACACTGACGCG 65
    |||||
Db 223 TTTGAGCTGATTGATGCTATAAAATGTGATCTTGAACGATATGCTGACACTGTTCT 282
    |||||

Qy 66 TGTGCGG 72
    |||||
Db 283 TGTGCTG 289

RESULT 4
BJ396832/c
LOCUS
DEFINITION
  BJ396832 Dictyostelium discoideum cDNA library, SF Dictyostelium
  discoideum cDNA clone dds44j03 5', mRNA sequence.

ORIGIN
  Query Match      37.3%; Score 30.2; DB 6; Length 449;
  Best Local Similarity 65.7%; Pred. No. 51;
  Matches 44; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 6 TTTGAAAAGATTGAGGCTGTGTTTGTGATGTGATTGTTGCTGACACTGACGCG 65
    |||||
Db 223 TTTGAGCTGATTGATGCTATAAAATGTGATCTTGAACGATATGCTGACACTGTTCT 282
    |||||

Qy 66 TGTGCGG 72
    |||||
Db 283 TGTGCTG 289

RESULT 4
BJ396832
LOCUS
DEFINITION
  BJ396832 Dictyostelium discoideum cDNA library, SF Dictyostelium
  discoideum cDNA clone dds44j03 5', mRNA sequence.
  
```

```

EST.
Eschscholzia californica (California poppy)
Eschscholzia californica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
Papaveraceae; Eschscholziaceae; Eschscholzia.
1 (bases 1 to 449)
dePamphilis,C., Carlson,J., Ma,H., Tanksley,S., Field,D.,
Leebens-Mack,J., Arrington,J., Zahn,L.M., Kong,H., Ilut,D.,
Druckemiller,M., Landherr,L., Hu,X., Plock,S., Wall,K.,
Chioorean,S., Albert,V., Doyle,J., Frohlich,M., Miller,W.,
Oppenheimer,D., Soltis,D., Soltis,P. and Theissen,G.
Generation of ESTs from early flower buds of Eschscholzia
californica (2002)
Unpublished
Contact: Claude dePamphilis or James Leebens-Mack
Mueller Laboratory
Penn State University
208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn
State University, University Park, PA 16802, USA
Tel: 814 863 6413
Fax: 814 865 9131
Email: cwd3@psu.edu or jhl10@psu.edu
The sequence provided is trimmed of vector and low quality regions.
Full sequence and original trace file are available from the Plant
Genome Network website (http://pgn.cornell.edu)
Plate: eca01-23msl row: g column: 10
Seq primer: M13F.
Location/Qualifiers
  1..449
  /organism="Eschscholzia californica"
  /mol_type="mRNA"
  /cultivar="Aurantia Orange"
  /db_xref="taxon:3467"
  /clone="eca01-23msl-g10"
  /tissue_type="flower buds <= 2.5mm"
  /dev_stage="millimeter buds"
  /lab_host="SOLR"
  /clone_lib="Eca01"
  /note="Vector: pBluescript SK (+/-); Site 1: EcoRI;
  Site 2: XhoI; Plants were grown in greenhouse at Penn
  State from commercially available seeds. Only floral buds
  with diameter of 2.5 mm or less were collected. This is a
  directionally cloned, non-normalized library. Avg. insert
  length: 1702; Primers: M13F and M13R; Antibiotic: 50 ug/ml
  Ampicillin; Primary Titer: 7E6 pfu total; Amplified Titer:
  1.8E11 pfu/ml; Mass Excised Titer: 5.6E8 total; This
  library has been generated by the Floral Genome Project
  Consortium for their assistance. The Floral Genome Project
  is funded by NSF's Plant Genome Research Program
  (DBI-0115684). More information about the project can be
  obtained at http://fsg.bio.psu.edu"

ORIGIN
  Query Match      37.3%; Score 30.2; DB 6; Length 449;
  Best Local Similarity 65.7%; Pred. No. 51;
  Matches 44; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 6 TTTGAAAAGATTGAGGCTGTGTTTGTGATGTGATTGTTGCTGACACTGACGCG 65
    |||||
Db 223 TTTGAGCTGATTGATGCTATAAAATGTGATCTTGAACGATATGCTGACACTGTTCT 282
    |||||

Qy 66 TGTGCGG 72
    |||||
Db 283 TGTGCTG 289

RESULT 4
BJ396832/c
LOCUS
DEFINITION
  BJ396832 Dictyostelium discoideum cDNA library, SF Dictyostelium
  discoideum cDNA clone dds44j03 5', mRNA sequence.
  
```

[illegible]

```

REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Okawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
MEDLINE 11076861
REFERENCE
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 11076861
REFERENCE
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirakawa, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
source
misc_feature 1..1533
/db_xref="taxon:10090"
/clone="9530079L20"
/sex="male"
/tissue_type="urinary bladder"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
/notes="unknown EST (GB|BE310219, evidence: BLASTN, 97%,
match=555)"
ORIGIN
Query Match 36.8%; Score 29.6; DB 3; Length 1533;
Best Local Similarity 70.2%; Pred. No. 85;
Matches 40; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
Qy 16 TTGAGGCTGTTTGTGATGCTGATTTGGATTGTGCTGACACTGACGGCTGTCGCG 72
Db 397 TTAATTGTTGTTTGTGATGCTGATTTGGATTGTGCTGACACTGACGGCTGTCGCG 453
RESULT 7
AV609193 684 bp mRNA linear EST 28-NOV-2001
LOCUS AV609193 Bos taurus lung fetus Bos taurus cDNA clone E1J021D12 5',
DEFINITION mRNA sequence.
ACCESSION AV609193.1 GI:9744863
VERSION AV609193
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE 1 (bases 1 to 684)
AUTHORS Takasuga, A., Hirotsune, S., Itoh, R., Jitohzono, A., Suzuki, H., Aso, H.
and Sugimoto, Y.
TITLE Establishment of a high throughput EST sequencing system using
poly(A) tail-removed cDNA libraries and determination of 36,000
bovine ESTs
JOURNAL Nucleic Acids Res. 29 (22), E108 (2001)
MEDLINE 21570554
PUBMED 11713328
COMMENT Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@coo.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
FEATURES
source
misc_feature 1..684
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="E1J021D12"
/tissue_type="lung"
/dev_stage="fetus"
/lab_host="DH10B"
/clone_lib="Bos taurus lung fetus"
/notes="Vector: pZ11; Site 1: SalI; Site 2: NotI; Poly A
was deleted from a NotI site"
ORIGIN
Query Match 36.5%; Score 29.6; DB 1; Length 684;
Best Local Similarity 64.7%; Pred. No. 85;
Matches 44; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 3 GAATTTGAAGATTGAGGCTGTTTGTGATGCTGATTTGGATTGTGCTGACACTGAC 62
Db 141 GAATTTGAAGATTGAGGCTGTTTGTGATGCTGATTTGGATTGTGCTGACACTGCTT 200

```

```

QY      63  GGCTGTGC 70
      |||||
Db      201  GGGTGTTC 208

RESULT 8
BX083182
LOCUS
DEFINITION
BX083182 748 bp mRNA linear EST 10-MAY-2004
(cac) AgENAE Rainbow trout normalized multi-tissues library
sequence.
ACCESSION
BX083182
VERSION
BX083182.2 GI:42607344
KEYWORDS
EST.
SOURCE
Oncorhynchus mykiss (rainbow trout)
ORGANISM
Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE
1 (bases 1 to 748)
Govoroun,M., Guiguen,Y. and Le Gac,F.
Construction and primary characterization of normalized cDNA
libraries in rainbow trout, Oncorhynchus mykiss
Unpublished (2003)
JOURNAL
On Jan 14, 2003 this sequence version replaced gi:27746599.
COMMENT
Contact: Guiguen Y
INRA - SCRIBE
Campus de Beaulieu, RENNES cedex, 35042, France
Tel: 02.23.48.50.09
Fax: 02.23.48.50.20
Email: Yann.Guiguen@beaulieu.rennes.inra.fr
Sequence cleaned of vector, adaptor and repetitions. Contact us
at sigenastupport@jouy.inra.fr to obtain the chromatogram of this
sequence.
Plate: 0004 row: f column: 24
Seq primer: T3.
FEATURES
source
Location/Qualifiers
1..748
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/clone="tcac0004c.f.24"
/tissue_type="adipose tissue, blood, brain,
differentiating gonads, gills, interrenal, intestine,
kidney, liver, muscle, ovary, pituitary, testis"
/dev_stage="from embryos to adults"
/lab_host="DH10B"
/clone_lib="AGENAE Rainbow trout normalized multi-tissues
library (tcac)"
/note="vector: pT73D-pac; Clone distribution : AGENAE
Resource centre: Francois PIUMI,
Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et
Etude du genome (LRBG), Domaine de Vilvert, 78352,
Jouy-en-Josas cedex, France, +33 (0) 1.34.65.28.02, +33
(0) 1.34.65.22.73"
ORIGIN
Query Match 36.5%; Score 29.6; DB 5; Length 748;
Best Local Similarity 64.7%; Pred. No. 86;
Matches 44; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY      2  TGAATTGAAAGATTGAGGCTGTTTGTGTGATGTTATGATTTGCTGCACTGA 61
      |||||
Db      245  TGCACAGGGAAGCGGCGAGCTGTCGTTGGTGATGTTGGGTGGGTGTTGTTGGCCTGA 304
      |||||

QY      62  CGGCTGTG 69
      |||||
Db      305  CCTGTGTG 312
      |||||

RESULT 9
BE052262
LOCUS
DEFINITION
GA_Ea0035L23f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboreum cDNA clone GA_Ea0035L23f, mRNA sequence.
ACCESSION
BE052262
VERSION
BE052262.2 GI:13246906
KEYWORDS
EST.
SOURCE
Gossypium arboreum
ORGANISM
Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE
1 (bases 1 to 804)
Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J.,
Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
JOURNAL
On Jun 8, 2000 this sequence version replaced gi:8379318.
COMMENT
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCATATAGGG
High quality sequence start: 2
High quality sequence stop: 475.
High quality sequence start: 475.
High quality sequence stop: 475.
FEATURES
source
Location/Qualifiers
1..804
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0035L23f"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
ORIGIN
Query Match 36.5%; Score 29.6; DB 2; Length 804;
Best Local Similarity 73.1%; Pred. No. 88;
Matches 38; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      3  GAATTTGAAAGATTGAGGCTGTTTGTGTGATGTTATGATTTGCTGCTG 54
      |||||
Db      474  GACTTTGATAAGCATGAGGATCTTGTGAGTGAATGTATAGGAGTTGAGCGG 525
      |||||

RESULT 10
CNS06ROG
LOCUS
DEFINITION
CNS06ROG 1050 bp DNA linear GSS 05-JUL-2001
T7 end of clone AM0AA016B08 of library AM0AA from strain CLIB 89 of
Yarrowia lipolytica, genomic survey sequence.
ACCESSION
AL412214
VERSION
AL412214.1 GI:12182513
KEYWORDS
GSS.
SOURCE
Yarrowia lipolytica
ORGANISM
Yarrowia lipolytica
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Dipodascaceae; Yarrowia.
REFERENCE
1 (bases 1 to 1050)
Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Boltelin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Malbertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissenbach,J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
FEBS Lett. 487 (1), 3-12 (2000)
JOURNAL

```



```

Plate: 0018 row: d column: 16
Seq primer: M13R. Location/Qualifiers
1. .498
  /organism="Oncorhynchus mykiss"
  /mol_type="mRNA"
  /db_xref="taxon:8022"
  /clone="tcbl0018c.d.16"
  /tissue_type="testis"
  /lab_host="DH10B"
  /clone_lib="AGENAE Rainbow trout normalized testis library (tcbl)"
  /note="Vector: pT7T3D-pac; Clone distribution : AGENAE Resource centre: Francois PIUMI, Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et Etude du genome (LREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex, FRANCE, +33 (0) 1.34.65.28.02, +33 (0) 1.34.65.22.73"

ORIGIN
Query Match 36.0%; Score 29.2; DB 7; Length 498;
Best Local Similarity 62.2%; Pred. No. 1.1e+02;
Matches 46; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1 ATGAATTTGAAAGATTGAGGCTGTTGTTGTGATGCTGATTGCTGACACTG 60
    |||||
Db 378 ATGAAGAGGAGAGACTGAGGCTATGCGCTCCGCTGTGCTTACATTGATGACACCTG 437
    |||||

QY 61 ACGGCTGTGCCGCG 74
    |||||
Db 438 CCAACCGTGACCGC 451

RESULT 13
CA381308 623 bp mRNA linear EST 06-NOV-2002
LOCUS 660794 NCCWA 1RT Oncorhynchus mykiss cDNA clone 1RT55A08_C_A04 5',
DEFINITION mRNA sequence.
ACCESSION CA381308
VERSION CA381308.1 GI:24703798
KEYWORDS EST.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Euteleostei;
  Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 623)
Rexroad, C.E. 3rd, Lee, Y., Keele, J.W., Karamycheva, S., Brown, G.,
Koop, B., Gahr, S.A., Palti, Y. and Quackenbush, J.
Sequence analysis of a rainbow trout cDNA library and creation of a
gene index
Cytogenet. Genome Res. 102 (1-4), 347-354 (2003)
COMMENT Contact: Rexroad CE
USDA, ARS, National Center for Cool and Cold Water Aquaculture
11876 Leetown Road, Kearneysville, WV 25430, USA
Tel: 304 724 8340 x2129
Fax: 304 725 0351
Email: crexroad@nccwa.ars.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified by
cross_match v0.990329
Seq primer: ACGGATACAAATTCACAGGA.
Location/Qualifiers
1. .623
  /organism="Oncorhynchus mykiss"
  /mol_type="mRNA"
  /db_xref="taxon:8022"
  /clone="1RT55A08_C_A04"
  /tissue_type="pooled"
  /lab_host="DH10B"
  /clone_lib="NCCWA 1RT"
  /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
  Library made from pooled tissue from brain, gill, liver,

```

```

spleen, muscle, and kidney."

ORIGIN
Query Match 36.0%; Score 29.2; DB 6; Length 623;
Best Local Similarity 62.2%; Pred. No. 1.1e+02;
Matches 46; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1 ATGAATTTGAAAGATTGAGGCTGTTGTTGTGATGCTGATTGCTGACACTG 60
    |||||
Db 479 ATGAAGAGGAGAGACTGAGGCTATGCGCTGTGCTTACATTGATGACACCTG 538
    |||||

QY 61 ACGGCTGTGCCGCG 74
    |||||
Db 539 CCAACCGTGACCGC 552

RESULT 14
BX859701 780 bp mRNA linear EST 21-MAY-2004
LOCUS BX859701 AGENAE Rainbow trout normalized testis library (tcbl)
DEFINITION Oncorhynchus mykiss cDNA clone tcbl0012d.c.19 5prim, mRNA sequence.
ACCESSION BX859701
VERSION BX859701.2 GI:42756450
KEYWORDS EST.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Euteleostei;
  Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 780)
Govoroun, M., Guiguen, Y. and Le Gac, F.
Construction and primary characterization of normalized cDNA
libraries in rainbow trout, Oncorhynchus mykiss
Unpublished (2003)
COMMENT On Dec 16, 2003 this sequence version replaced gi:39957467.
Contact: Guiguen Y
INRA - SCRIBE
Campus de beaulieu, RENNES cedex, 35042, France
Tel: 02.23.48.50.09
Fax: 02.23.48.50.20
Email: Yann.Guiguen@beaulieu.rennes.inra.fr
Sequence cleaned of vector, adaptor and repetitions. Contact us
at signenasupport@jouy.inra.fr to obtain the chromatogram of this
sequence.
Plate: 0012 row: c column: 19
Seq primer: M13R. Location/Qualifiers
1. .780
  /organism="Oncorhynchus mykiss"
  /mol_type="mRNA"
  /db_xref="taxon:8022"
  /clone="tcbl0012d.c.19"
  /tissue_type="testis"
  /lab_host="DH10B"
  /clone_lib="AGENAE Rainbow trout normalized testis library (tcbl)"
  /note="Vector: pT7T3D-pac; Clone distribution : AGENAE Resource centre: Francois PIUMI, Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et Etude du genome (LREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex, FRANCE, +33 (0) 1.34.65.28.02, +33 (0) 1.34.65.22.73"

ORIGIN
Query Match 36.0%; Score 29.2; DB 5; Length 780;
Best Local Similarity 62.2%; Pred. No. 1.2e+02;
Matches 46; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1 ATGAATTTGAAAGATTGAGGCTGTTGTTGTGATGCTGATTGCTGACACTG 60
    |||||
Db 308 ATGAAGAGGAGAGACTGAGGCTATGCGCTCCGCTGTGCTTACATTGATGACACCTG 367
    |||||

QY 61 ACGGCTGTGCCGCG 74
    |||||

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 10, 2004, 21:26:43 ; Search time 423.895 Seconds  
(without alignments)  
8421.106 Million cell updates/sec

Title: US-09-909-207-1  
Perfect score: 663  
Sequence: 1 CAATCGTCACCGACAATTC.....TAATCTTGATATAAAACAAT 663

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3611042 seqs, 2692057975 residues  
Total number of hits satisfying chosen parameters: 7222084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:  
4: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:  
5: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
6: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:  
7: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:  
8: /cgn2\_6/prodata/1/pubpna/US09A\_PUBCOMB.seq:  
9: /cgn2\_6/prodata/1/pubpna/US09B\_PUBCOMB.seq:  
10: /cgn2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:  
11: /cgn2\_6/prodata/1/pubpna/US09D\_PUBCOMB.seq:  
12: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq:  
13: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:  
14: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:  
15: /cgn2\_6/prodata/1/pubpna/US10D\_PUBCOMB.seq:  
16: /cgn2\_6/prodata/1/pubpna/US10E\_PUBCOMB.seq:  
17: /cgn2\_6/prodata/1/pubpna/US10F\_PUBCOMB.seq:  
18: /cgn2\_6/prodata/1/pubpna/US10G\_PUBCOMB.seq:  
19: /cgn2\_6/prodata/1/pubpna/US10H\_PUBCOMB.seq:  
20: /cgn2\_6/prodata/1/pubpna/US10I\_PUBCOMB.seq:  
21: /cgn2\_6/prodata/1/pubpna/US10J\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	663	100.0	663	9	US-09-909-207-1
2	663	100.0	663	9	US-09-909-207-2
3	663	100.0	744	9	US-09-909-207-4
4	663	100.0	744	9	US-09-909-207-5
5	663	100.0	1513	9	US-09-909-207-10
6	663	100.0	1513	9	US-09-909-207-11
7	119.4	18.0	1375	9	US-09-770-621-1
8	119.4	18.0	1375	15	US-10-286-993-1
9	115.2	17.4	596	15	US-10-307-441-39
10	111.4	16.8	942	14	US-10-213-990-71
11	110	16.6	1002	14	US-10-213-990-70
12	108.8	16.4	705	14	US-10-213-990-68

13	106	16.0	2898	15	US-10-299-393-1	Sequence 1, Appli
14	104.8	15.8	712	14	US-10-213-990-64	Sequence 64, Appl
15	103.6	15.6	666	14	US-10-213-990-65	Sequence 65, Appl
16	96.8	14.6	739	14	US-10-213-990-67	Sequence 67, Appl
17	86.8	13.1	678	10	US-09-803-454-3	Sequence 3, Appli
18	84.2	12.7	1023	16	US-10-244-596-6	Sequence 6, Appli
19	84.2	12.7	1027	16	US-10-244-596-2	Sequence 2, Appli
20	83.2	12.5	645	15	US-10-237-386-10	Sequence 10, Appl
21	83.2	12.5	657	15	US-10-237-386-11	Sequence 11, Appl
22	82.6	12.5	1011	16	US-10-244-596-12	Sequence 12, Appl
23	82.6	12.5	1011	16	US-10-244-596-13	Sequence 13, Appl
24	82.6	12.5	1023	16	US-10-244-596-5	Sequence 5, Appli
25	82.6	12.5	1023	16	US-10-244-596-7	Sequence 7, Appli
26	82.6	12.5	1023	16	US-10-244-596-8	Sequence 8, Appli
27	82.6	12.5	1023	16	US-10-244-596-9	Sequence 9, Appli
28	82.6	12.5	1023	16	US-10-244-596-3	Sequence 3, Appli
29	81	12.2	1023	16	US-10-244-596-10	Sequence 10, Appl
30	79.4	12.0	1023	16	US-10-244-596-4	Sequence 4, Appli
31	79.4	12.0	1023	16	US-09-790-070A-8	Sequence 8, Appli
32	77.8	11.7	1014	16	US-10-244-596-1	Sequence 1, Appli
33	77.8	11.7	749	18	US-10-425-115-82922	Sequence 82922, A
34	77.2	11.6	818	18	US-10-425-115-37173	Sequence 37173, A
35	75.8	11.4	850	18	US-10-425-115-177283	Sequence 177283,
36	73.8	11.1	588	15	US-10-237-386-9	Sequence 9, Appli
37	72.4	10.9	983	9	US-09-467-368-1	Sequence 1, Appli
38	72.4	10.9	2054	15	US-10-419-969-5	Sequence 5, Appli
39	54.2	8.2	643	15	US-10-340-860A-39	Sequence 39, Appl
40	50.4	7.6	74	15	US-10-307-441-24	Sequence 24, Appl
41	41	6.2	557	9	US-09-970-616-1	Sequence 1, Appli
42	40	6.0	642	11	US-09-149-310-31	Sequence 31, Appl
43	40	6.0	6123	15	US-10-311-455-793	Sequence 793, App
C 44	38	5.7	1792	15	US-10-369-493-29835	Sequence 29835, A
C 45	37.8	5.7				

ALIGNMENTS

RESULT 1  
US-09-909-207-1  
; Sequence 1, Application US/09090207  
; Patent No. US20020145181A1  
; GENERAL INFORMATION:  
; APPLICANT: ANDREE LAHAYE  
; ERIC DE RUYA  
; PATRICE LEDOUX  
; RENE DETROZ

TITLE OF INVENTION: Xylanase, microorganisms produced it,  
DNA molecule, processes for preparation of this xylanase  
and uses thereof

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE

STREET: 2000 K St., N.W., Suite 200

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/09/909,207

APPLICATION NUMBER: 19-Jul-2001

FILING DATE: 19-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/470,953

FILING DATE: 06-JUNE-1995

ATTORNEY/AGENT INFORMATION:

NAME: Wilhem F. Gadiano, Esq.

REGISTRATION NUMBER: 37,136

Patent No. US20020115181A1  
GENERAL INFORMATION:  
APPLICANT: ANDREE LAHAYE  
ERIC DE BUYL  
PIERRE LEDOUX  
RENE DETROZ  
TITLE OF INVENTION: Xylanase, microorganisms produced it,  
DNA molecule, processes for preparation of this xylanase  
and uses thereof  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
STREET: 2000 K St., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.

ZIF: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/09/909,207  
APPLICATION NUMBER: US/09/909,207  
FILING DATE: 19-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 05/470,953  
FILING DATE: 06-JUNE-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilhlem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-40  
TELECOMMUNICATION INFORMATION:

181 CAACAAGTTGGTAACATGTCATATAAATCTCGAGGCCAACCTTCCACCAAATGGTAATGCG 240

Db 181 CAACAAAGTTGGTAAACATGTCATTAACATACGAGGCAACCTTCCACCAAAATGGTAATCGG 240  
QY 241 TATTTATGCGTCTATGTTGGAGCTGTTGACCCCTCTTGCGAATATATATATGTCACAGT 300  
Db 241 TATTTATGCGTCTATGTTGGAGCTGTTGACCCCTCTTGCGAATATATATATGTCACAGT 300  
QY 301 TGGGGCAACTGCGCTCCACGAGGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 360  
Db 301 TGGGGCAACTGCGCTCCACGAGGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 360  
QY 361 ACATATGATATACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATGCCACA 420  
Db 361 ACATATGATATACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATGCCACA 420  
QY 421 TTTAAACAATATGAGTGTTCGAGATCGAAACGACGAGTGGCAAGATTTCTGTGAGC 480  
Db 421 TTTAAACAATATGAGTGTTCGAGATCGAAACGACGAGTGGCAAGATTTCTGTGAGC 480  
QY 481 AACCCTTTAGAGCGTGGGAAACCTTAGGATGAATATGCGGAAATGTATGAAGTCGCG 540  
Db 481 AACCCTTTAGAGCGTGGGAAACCTTAGGATGAATATGCGGAAATGTATGAAGTCGCG 540  
QY 541 CTCTACTGTAGAAGCTTATCAAGTACGGAAGTGTAAATGTATATAGCAATACACTAAGA 600  
Db 541 CTCTACTGTAGAAGCTTATCAAGTACGGAAGTGTAAATGTATATAGCAATACACTAAGA 600  
QY 601 ATTAACGGTAACCTCTCTCAACTATATAGTATGACGAGACGATTAATTTGGATAAAAC 660  
Db 601 ATTAACGGTAACCTCTCTCAACTATATAGTATGACGAGACGATTAATTTGGATAAAAC 660  
QY 661 AAT 663  
Db 661 AAT 663

RESULT 3

US-09-909-207-4  
; Sequence 4, Application US/0909207  
; Patent No. US20020115181A1  
; GENERAL INFORMATION:  
; APPLICANT: ANDREE LAHAYE  
; ERIC DE BUYL  
; PIERRE LEDOUX  
; RENE DETROZ  
; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
; DNA molecule, processes for preparation of this xylanase  
; and uses thereof  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: WILLIAN BRINKS HOPER GILSON & LIONE  
; STREET: 2000 K St., N.W., Suite 200  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/909,207  
; FILING DATE: 19-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/470,953  
; FILING DATE: 06-JUNE-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wilhem F. Gadiano, Esq.  
; REGISTRATION NUMBER: 37,136  
; REFERENCE/DOCKET NUMBER: 4121-40  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-429-0625

TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 744 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; STRAIN: Bacillus  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-909-207-4  
Query Match 100.0%; Score 663; DB 9; Length 744;  
Best Local Similarity 100.0%; Pred. No. 1.6e-183;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAAATCGTCACGCAAAATTCATTCGGCAACCAACGATGGCTATGATATGAATTTGGAAA 60  
Db 82 CAAATCGTCACGCAAAATTCATTCGGCAACCAACGATGGCTATGATATGAATTTGGAAA 141  
QY 61 GATAGCGGTGGCTCTGGGACATGATTTCTCAATCATGGGGTACCTTCAGTCCCAATCG 120  
Db 142 GATAGCGGTGGCTCTGGGACATGATTTCTCAATCATGGGGTACCTTCAGTCCCAATCG 201  
QY 121 AACAAATGTTAAACAATATTTCCGTAAAGGTAAATAATTCATGAAACACACACAC 180  
Db 202 AACAAATGTTAAACAATATTTCCGTAAAGGTAAATAATTCATGAAACACACACAC 261  
QY 181 CAACAAGTTGTAACATGTCCTCAATAAATACGAGGCAACTTCCAAACCAATGGTAATGG 240  
Db 262 CAACAAGTTGTAACATGTCCTCAATAAATACGAGGCAACTTCCAAACCAATGGTAATGG 321  
QY 241 TATTTATGCGTCTATGTTGGAGTCTGTGACCTCTCTGTCGATATATATATGTCGACAGT 300  
Db 322 TATTTATGCGTCTATGTTGGAGTCTGTGACCTCTCTGTCGATATATATATGTCGACAGT 381  
QY 301 TGGGGCAACTGCGCTCCACGAGGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 360  
Db 382 TGGGGCAACTGCGCTCCACGAGGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 441  
QY 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTCACACA 420  
Db 442 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTCACACA 501  
QY 421 TTTAAACAATATGAGTGTTCGAGATCGAAACGACGAGTGGCAAGATTTCTGTGAGC 480  
Db 502 TTTAAACAATATGAGTGTTCGAGATCGAAACGACGAGTGGCAAGATTTCTGTGAGC 561  
QY 481 AACCCTTTAGAGCGTGGGAAACCTTAGGATGAATATGCGGAAATGTATGAAGTCGCG 540  
Db 562 AACCCTTTAGAGCGTGGGAAACCTTAGGATGAATATGCGGAAATGTATGAAGTCGCG 621  
QY 541 CTCTACTGTAGAAGCTTATCAAGTACGGAAGTGTAAATGTATATAGCAATACACTAAGA 600  
Db 622 CTCTACTGTAGAAGCTTATCAAGTACGGAAGTGTAAATGTATATAGCAATACACTAAGA 681  
QY 601 ATTAACGGTAACCTCTCTCAACTATATAGTATGACGAGACGATTAATTTGGATAAAAC 660  
Db 682 ATTAACGGTAACCTCTCTCAACTATATAGTATGACGAGACGATTAATTTGGATAAAAC 741  
QY 661 AAT 663  
Db 742 AAT 744

RESULT 4

US-09-909-207-5  
; Sequence 5, Application US/0909207  
; Patent No. US20020115181A1  
; GENERAL INFORMATION:  
; APPLICANT: ANDREE LAHAYE

ERIC DE BUYL  
PIERRE LEDOUX  
RENE DETROZ  
TITLE OF INVENTION: Xylanase, microorganisms produced it,  
DNA molecule, processes for preparation of this xylanase  
and uses thereof  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
STREET: 2000 K St., N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/909,207  
FILING DATE: 19-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/470,953  
FILING DATE: 06-JUNE-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilhem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-40  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 744 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: Bacillus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..744  
NAME/KEY: mat\_peptide  
LOCATION: 82..744  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 1..81  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-909-207-5  
Query Match 100.0%; Score 663; DB 9; Length 744;  
Best Local Similarity 100.0%; Pred. No. 1.6e-183; Indels 0; Gaps 0;  
Matches 663; Conservative 0; Mismatches 0;  
QY 1 CAAATCGTCACCGACAATTCATTTGGCAACACGATGGCTATGATTGAAATTTGGAAA 60  
DB 82 CAAATCGTCACCGACAATTCATTTGGCAACACGATGGCTATGATTGAAATTTGGAAA 141  
QY 61 GATAGCGGTGGCTCTGGGCAATGATCTCAATCATGCGGGTACGTTCAATGCGCAATGG 120  
DB 142 GATAGCGGTGGCTCTGGGCAATGATCTCAATCATGCGGGTACGTTCAATGCGCAATGG 201  
QY 121 AACAAATGTTAAACAATATTTTCGTTAAAGTAAATAATTCATGAAACACAAACACAC 180  
DB 202 AACAAATGTTAAACAATATTTTCGTTAAAGTAAATAATTCATGAAACACAAACACAC 261  
QY 181 CAAACAAGTTGGTAACATGTCCTAAATACGAGGCAACTTCCAAACCAATGGTAATGCG 240

DB 262 CAAACAAGTTGGTAACATGTCCTAAATACGAGGCAACTTCCAAACCAATGGTAATGCG 321  
QY 241 TATTTATGCGTCTATGTTGGAGCTGTGACCCCTTGTGCGAATATATATTTGTCGACAGT 300  
DB 322 TATTTATGCGTCTATGTTGGAGCTGTGACCCCTTGTGCGAATATATATTTGTCGACAGT 381  
QY 301 TGGGGCAACTGGCGTCCACGAGGACAAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 360  
DB 382 TGGGGCAACTGGCGTCCACGAGGACAAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 441  
QY 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTCGCACA 420  
DB 442 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTCGCACA 501  
QY 421 TTTTAAACAATATTGGAGTGTTCGAAGATCGAAGCGACGAGTGGCAGCATTTCTGTGTCAGC 480  
DB 502 TTTTAAACAATATTGGAGTGTTCGAAGATCGAAGCGACGAGTGGCAGCATTTCTGTGTCAGC 561  
QY 481 AACCACTTTTAGAGCGTGGGAAAACCTTAGGGATGAATATGGGAAAATGTATGAAGTGGCG 540  
DB 562 AACCACTTTTAGAGCGTGGGAAAACCTTAGGGATGAATATGGGAAAATGTATGAAGTGGCG 621  
QY 541 CTTACTGTAGAAGCTATCAAAAGTAGCGGAAGTCTTAATGTATATAGCAATACACTAAGA 600  
DB 622 CTTACTGTAGAAGCTATCAAAAGTAGCGGAAGTCTTAATGTATATAGCAATACACTAAGA 681  
QY 601 ATTTAACGGTAACCCCTCTCTCAACTATTAGTAATGACGAGAGCATAACTTTTGGATAAAAAC 660  
DB 682 ATTTAACGGTAACCCCTCTCTCAACTATTAGTAATGACGAGAGCATAACTTTTGGATAAAAAC 741  
QY 661 AAT 663  
DB 742 AAT 744  
RESULT 5  
US-09-909-207-10  
; Sequence 10, Application US/09909207  
; Patent No. US20020115181A1  
; GENERAL INFORMATION:  
; APPLICANT: ANDREE LAHAYE  
; ERIC DE BUYL  
; PIERRE LEDOUX  
; RENE DETROZ  
; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
; DNA molecule, processes for preparation of this xylanase  
; and uses thereof  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
; STREET: 2000 K St., N.W., Suite 200  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/909,207  
; FILING DATE: 19-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/470,953  
; FILING DATE: 06-JUNE-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wilhem F. Gadiano, Esq.  
; REGISTRATION NUMBER: 37,136  
; REFERENCE/DOCKET NUMBER: 4121-40  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-429-0625

TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1513 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: Bacillus  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-09-909-207-10

Query Match 100.0%; Score 663; DB 9; Length 1513;  
Best Local Similarity 100.0%; Pred. No. 2.2e-183;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAATCGTCACCGACAATTCCTATGGCAACCGATGGCTATGATTTATGAATTTTGGAAA 60  
DB 701 CAAATCGTCACCGACAATTCCTATGGCAACCGATGGCTATGATTTATGAATTTTGGAAA 760

QY 61 GATAGCGGTGGCTCTCGGACAATGATTTCAATCATGCGGTACGTTAGTCCCAATGG 120  
DB 761 GATAGCGGTGGCTCTCGGACAATGATTTCAATCATGCGGTACGTTAGTCCCAATGG 820

QY 121 AACATGTTAAACAATATTTTCCGTAAGGTAAAGTAAATTTCAATGAACACAAACACAC 180  
DB 821 AACATGTTAAACAATATTTTCCGTAAGGTAAAGTAAATTTCAATGAACACAAACACAC 880

QY 181 CAACAAGTTGGTAACATGTCATAAATCGGAGCCAACTTCCACCAATGTTATGCG 240  
DB 881 CAACAAGTTGGTAACATGTCATAAATCGGAGCCAACTTCCACCAATGTTATGCG 940

QY 241 TATTTATGCGCTATGTTGGACTGTTGACCTCTTGTGCAATATTTATTTGTCACAGT 300  
DB 941 TATTTATGCGCTATGTTGGACTGTTGACCTCTTGTGCAATATTTATTTGTCACAGT 1000

QY 301 TGGGGCAACTGGCGTCCACGAGGCAACGCCCTAAGGGACCACATCTGTTGATGAGGA 360  
DB 1001 TGGGGCAACTGGCGTCCACGAGGCAACGCCCTAAGGGACCACATCTGTTGATGAGGA 1060

QY 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGATGTCACA 420  
DB 1061 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGATGTCACA 1120

QY 421 TTTTAAACAATTTGGAGTGTTCGAAGATCGAAACGCAAGTGGCCAGATTTCTGTCAGC 480  
DB 1121 TTTTAAACAATTTGGAGTGTTCGAAGATCGAAACGCAAGTGGCCAGATTTCTGTCAGC 1180

QY 481 AACCACTTTAGAGGTGGGAAAATTTAGGATGAATATGGGAAAATGATGATGAGTCGCG 540  
DB 1181 AACCACTTTAGAGGTGGGAAAATTTAGGATGAATATGGGAAAATGATGATGAGTCGCG 1240

QY 541 CTCTACTGTAGAGCTATCAAGTAGCGGAAGTCTAATGATATATAGCAATACACTAAGA 600  
DB 1241 CTCTACTGTAGAGCTATCAAGTAGCGGAAGTCTAATGATATATAGCAATACACTAAGA 1300

QY 601 ATTAACCGGTAAACCTCTCTCAACTATTAGTAAATGACGAGAGCACTTTGGTAAACAC 660  
DB 1301 ATTAACCGGTAAACCTCTCTCAACTATTAGTAAATGACGAGAGCACTTTGGTAAACAC 1360

QY 661 AAT 663  
DB 1361 AAT 1363

RESULT 6  
US-09-909-207-11  
; Sequence 11, Application US/0909207  
; Patent No. US20020115181A1  
; GENERAL INFORMATION:  
; APPLICANT: ANDREE LAHAYE

ERIC DE BUYL  
PIERRE LEDOUX  
RENE DETROZ  
TITLE OF INVENTION: Xylanase, microorganisms produced it,  
DNA molecule, processes for preparation of this xylanase  
and uses thereof  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
STREET: 2000 K St., N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/909,207  
FILING DATE: 19-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/470,953  
FILING DATE: 06-JUNE-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilhem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-40  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1513 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: Bacillus  
NAME/KEY: CDS  
LOCATION: 620..1363  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 701..1363  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 620..700  
SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-09-909-207-11

Query Match 100.0%; Score 663; DB 9; Length 1513;  
Best Local Similarity 100.0%; Pred. No. 2.2e-183;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAATCGTCACCGACAATTCCTATGGCAACCGATGGCTATGATTTATGAATTTTGGAAA 60  
DB 701 CAAATCGTCACCGACAATTCCTATGGCAACCGATGGCTATGATTTATGAATTTTGGAAA 760

QY 61 GATAGCGGTGGCTCTCGGACAATGATTTCAATCATGCGGTACGTTAGTCCCAATGG 120  
DB 761 GATAGCGGTGGCTCTCGGACAATGATTTCAATCATGCGGTACGTTAGTCCCAATGG 820

QY 121 AACATGTTAAACAATATTTTCCGTAAGGTAAAGTAAATTTCAATGAACACAAACACAC 180  
DB 821 AACATGTTAAACAATATTTTCCGTAAGGTAAAGTAAATTTCAATGAACACAAACACAC 880

QY 181 CAACAAGTTGGTAACATGTCATAAATCGGAGCCAACTTCCACCAATGTTATGCG 240  
DB 881 CAACAAGTTGGTAACATGTCATAAATCGGAGCCAACTTCCACCAATGTTATGCG 240

```

881 CAACAGTTGGTAAACATGCTCCATAAACTACGAGGCAACCTTCCACCAATGGTATCG 940
241 TATTATGCGTCTATGGTTGGACTGTTGACCCCTTTGTCGAATATATATGTCGACAGT 300
941 TATTATGCGTCTATGGTTGGACTGTTGACCCCTTTGTCGAATATATATGTCGACAGT 1000
301 TGGGGCAACTGGCGTCCACGAGGCAACGCTTAAGGGGACCATCACTGTTGATGAGGA 360
1001 TGGGGCAACTGGCGTCCACGAGGCAACGCTTAAGGGGACCATCACTGTTGATGAGGA 1060
361 ACATATGATATACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTCACACA 420
1061 ACATATGATATACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTCACACA 1120
421 TTTAAACATATGAGGTTTCGAAGATCGAAACGACGAGTGGACCATTTCTGTCAGC 480
1121 TTTAAACATATGAGGTTTCGAAGATCGAAACGACGAGTGGACCATTTCTGTCAGC 1180
481 AACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATGGGAAATGTATGAAGTCGG 540
1181 AACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATGGGAAATGTATGAAGTCGG 1240
541 CTTACTGTAGAGGCTATCAAAAGTACGGAAGTGTATATATAGCAATACACTAAGA 600
1241 CTTACTGTAGAGGCTATCAAAAGTACGGAAGTGTATATATAGCAATACACTAAGA 1300
601 ATTAACGGTAACCTCTCTCAACTATTAGTATGACGAGGACATCACTTGGATAAAAAC 660
1301 ATTAACGGTAACCTCTCTCAACTATTAGTATGACGAGGACATCACTTGGATAAAAAC 1360
661 AAT 663
1361 AAT 1363

```

# RESULT 7

```

US-09-770-621-1
; Sequence 1, Application US/09770621
; Patent No. US20010024815A1
; GENERAL INFORMATION:
; APPLICANT: M ntyl, Arja
; APPLICANT: Vehmanper, Jari
; APPLICANT: Fagerstr m, Richard
; APPLICANT: Lantto, Raija
; APPLICANT: Paloheimo, Marja
; APPLICANT: Suominen, Pirkko
; APPLICANT: Lahtinen, Tarja
; TITLE OF INVENTION: Production and Secretion of Proteins of
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 New York Ave., N.W. Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/770,621
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/590,563
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/332,412
; FILING DATE: 31-OCT-1994
; CLASSIFICATION:

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/282,001
; FILING DATE: 29-JUL-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugalsky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 1050.0340003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1375 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: No. US20010024815A1 Relevant
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 303..1334
; US-09-770-621-1
; Query Match 18.0%; Score 119.4; DB 9; Length 1375;
; Best Local Similarity 57.1%; Pred. No. 3.8e-24;
; Matches 238; Conservative 0; Mismatches 176; Indels 3; Gaps 1;
QY 198 GTCATAACTACGAGGCAACCTTCCAAACCAATGGTAATGCGTATTTATGCGTCTATGG 257
Db 608 GACCGTGACCTTACAAACGCTCTTCAACCCGTCGGGTAAAGGCTACCTCAAGCTCTACGG 667
QY 258 TTGCACTGTTGACCCCTCTGTGCAATATATATGTCGACAGTTGGGCAACTGGCGTCC 317
Db 668 CTGACCAAGAACCCGCTCGTCGAGTACTACATCGTCGAGAGCTGGGACCTACCGGCC 727
QY 318 ACCAGGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGAACATATGATATCTACGA 377
Db 728 CACCGG--CACCTTACAAGGGGACCGTCACCACCGCGGGGAAACGTTACGACATCTACGA 784
QY 378 GACTCTTAGAGTCAATCAACCTCCATTAAGGGGATGCCACATTTAAACATATTGGAG 437
Db 785 GACCTGGCGGTACAAACGCGCGCTCCATCGAGGGGACCCGACCTTCCAGCAGTTCTGGAG 844
QY 438 TGTTGGAAGATCGAAACGACGAGTGGCAGCATTTCTGTCAGCAACACCTTTAGAGGGTG 497
Db 845 CGTCGGGACGAGAAAGGGACCGGACCAATCACCATCGGCAACCACTTCGACGCGTG 904
QY 498 GGAATACTTAGGGATGAATATGGGAAATGTATGAAGTCCGCTTACTGTAGAGGCTA 557
Db 905 GGCCCGCGCGGATGAACCTGGGCGACCACTACCATCATGGGACCGAGGCGTA 964
QY 558 TCAAGTAGCGGAGTGTCTAATGTTATATAGCAATACACTAAGATTAAACGTTACCC 614
Db 965 CCAGAGCAGCGGTAGTCTCCACCGGTCTCCATCAGCGGGTGGCAACCCCGGCAACCC 1021
;
; RESULT 8
; US-10-286-993-1
; Sequence 1, Application US/10286993
; Publication No. US20030148453A1
; GENERAL INFORMATION:
; APPLICANT: Mantlya, Arja
; APPLICANT: Paloheimo, Marja
; APPLICANT: Lantto, Raija
; APPLICANT: Fagerstrom, Richard
; APPLICANT: Lahtinen, Tarja
; APPLICANT: Suominen, Pirkko
; APPLICANT: Vehmanpera, Jari
; TITLE OF INVENTION: Production and Secretion of Proteins in Filamentous
; FILING DATE: 1716.034004
; FILE REFERENCE: 1716.034004
; CURRENT APPLICATION NUMBER: US/10/286,993
; CURRENT FILING DATE: 2002-08-13

```

```
; PRIOR APPLICATION NUMBER: US/09/120,804
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: PCT/FI97/00037
; PRIOR FILING DATE: 1997-01-24
; PRIOR APPLICATION NUMBER: US 08/590,563
; PRIOR FILING DATE: 1996-01-26
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1375
; TYPE: DNA
; ORGANISM: Actinomadura flexuosa (Strain: DSM43186)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (303)..(1337)
; OTHER INFORMATION: Product= AM35 xylanase
; US-10-286-993-1

Query Match      18.0%; Score 119.4; DB 15; Length 1375;
Best Local Similarity 57.1%; Pred. No. 3.8e-24;
Matches 238; Conservative 0; Mismatches 176; Indels 3; Gaps 1;

QY 198 GTCCATAAACTACGGAGCCAACTTCCAAACCAATGTAATGTAATTTATGCGTCTATGG 257
DB 608 GACCGTGACCTACAAACGCTCTCTTCAACCGTGGGTAAACGGCTACCTCAGCGCTTACGG 667

QY 258 TTGAGCTGTGACCTCTTGTGCAATATTATATTGTCGACAGTGTGGGGCAACTGGCGTCC 317
DB 668 GTGACCAAGAAACCGCTCGTGTGAGTACTACATCGTCGAGAGCGTGGGGCACTTACCGGC 727

QY 318 ACCGAGGCAACGCTTAAGGGGACCATCACTGTGTTGATGAGGAAACATATGATATCTACGA 377
DB 728 CACCGG---CACCTCAAGGGACCGCTCACCCAGCGGGGAGCTAGACATCTACGA 784

QY 378 GACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTTGCCACATTTTAAACAATATTGGAG 437
DB 785 GACTCGCGGTACAAACGCGCGTCCATCGAGGGCACCCGCGACCTTCCAGCAGTTCTGGAG 844

QY 438 TGTTGGAAGATCGAAACGACGAGTGGCGAGATTTCTGTGACGAAACCACTTTAGACGCTG 497
DB 845 CGTCCGGCAGCAGAGCGGACGACGAGCGCACCATCAGCATCGGCAACCACTTCGACGCTG 904

QY 498 GGAATACTAGGATGAATATGGGAAATGTATGATGATGATGATGATGATGATGATGATGATG 557
DB 905 GGCCCGCGCGGATGAACCTGGGCGACGACGACTACAGATCATGGGACCGGGCTA 964

QY 558 TCAAAGTAGCGGAAGTGTCTAATGTATATGCAATACACTAAGAAATTAACCGTAAACCC 614
DB 965 CCAGACGAGCGGTAGCTTCCACCGTCTCCATCAGGAGGGTGGCAACCCCGCAACCC 1021

RESULT 9
US-10-307-441-39
; Sequence 39, Application US/10307441
; Publication No. US20030166236A1
; GENERAL INFORMATION:
; APPLICANT: SUNG, Wing L.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Modified Xylanases Exhibiting Increased Thermophilicity
; TITLE OF INVENTION: and Alkalophilicity
; FILE REFERENCE: 027367-5006US
; CURRENT APPLICATION NUMBER: US/10/307,441
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: PCT/CA01/00769
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/213,803
; PRIOR FILING DATE: 2000-05-31
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 596
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Trx
; US-10-307-441-39

Query Match      17.4%; Score 115.2; DB 15; Length 596;
Best Local Similarity 57.7%; Pred. No. 4.3e-23;
Matches 226; Conservative 0; Mismatches 163; Indels 3; Gaps 1;

QY 193 AACATGTCCATAAACTACGGAGCCAACTTCCAAACCAATGGTAATGCGTATTATGCGTC 252
DB 191 AATAAGGTGATCAACTTCTTGATCTTATAATCCGAATGGGAATTCATACTTAAGCGTC 250

QY 253 TATGTTGGAGCTTTGACCCCTTCTGCAATATTATATTGTCGACAGTTGGGGCACTGG 312
DB 251 TATGCTGCTCTAGAAACCCCACTGATTGAATATTATGTCGAAAAATTCGGTACCTTAC 310

QY 313 CGTCC---ACCAGGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGAACATATGAT 369
DB 311 AATCCGAGTACCGGCGCCACAAAATTAGCGGAAGTCACTAGTGTGATCCGTATATGAT 370

QY 370 ATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAAGGGGATTTGCCACATTTAAACAA 429
DB 371 ATCTACCGTACCAACGCGTTAATCAGCATCGATCATTTGGAACCGCCACCTTTTATCAG 430

QY 430 TATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGCAATTTCTGTGACGAAACCACTTT 489
DB 431 TACTGGAGTGTTAGACGTAATCATCGGAGCTCCGGTTCGGTTAAATACTCGGAATCACTTT 490

QY 490 AGAGCGTGGGAAACTTAGGGATGAATATGGGAAATGTATGAAGTCGCGCTTACTGTA 549
DB 491 AATGATCGGCAACGCAAGGGTTAACCCCTAGGTACAATGGATTAATCAAAATCGTAGCGGTG 550

QY 550 GAAGCTATCAAAAGTAGCGGAAGTGTCTAATGT 581
DB 551 GAAGGCTACTTCTCGAGTGTTCGCGTAGTAT 582

RESULT 10
US-10-213-990-71
; Sequence 71, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Storms, Reg
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
; TITLE OF INVENTION: ENZYMES AND METHODS OF USE
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Aspergillus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(942)
; US-10-213-990-71

Query Match      16.8%; Score 111.4; DB 14; Length 942;
Best Local Similarity 57.0%; Pred. No. 6.9e-22;
Matches 224; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

QY 182 AACAGTTGGTAACATGTCCATAAACTACGGAGCCAACTTTCCAAACCAATATGTAATCGGT 241
DB 257 ATCCAGGAGTGACCATGACATTAATCACTTCTCTGGCAGCTTCAATCTTCGGAATGCTT 316

QY 242 ATTTATGCGTCTATGTTGACCTTTGTGACCTTTGTGCAATATTATATTGTCGACAGTT 301
DB 317 ACCTGTCGGTGTATGGATGAGTACCAACCCCTAGTCGAATACTACTACATCTCTCGAAGT 376
```

```
QY 302 GGGCA---ACTGGCGTCCACGAGCAACGCTTAAGGGACCATCACTGTTGATGAG 358
Db 377 ATGGCAGTTACAACTCCTGGCTCGGCGATGACGCAAGGGCACCGCTCACCGAGGATGAT 436
QY 359 GAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAAGGGGATGCGCA 418
Db 437 CCACCTAGACATCTATGACGACCAACAGGTCAACGAGCTTCGATCTGGCAGCGGCA 496
QY 419 CATTAAACAATATTGGAGTGTGGAAGATCGAAACGACGAGTGGCAGCATTTCTGTGCA 478
Db 497 CCTTCAACCAATATCTGGTCCATCCGCCAAAACAAGCGATCCAGCGGCACAGTCAACACCG 556
QY 479 GCAACCACTTTAGAGCGTGGGAAACTTAGGATGAATATCGGGAAATGTATGAAGTCG 538
Db 557 CGAATCACTCAAGGCCTGGGCTAGTCTGGGGATGAACCTGGGTACCCATTAATCATCAGA 616
QY 539 CGCTTACTGTAGAAGGCTATCAAAAGTAGCGGAA 571
Db 617 TTGTTTCCACTGAGGGATATGAGACGCGGTA 649

RESULT 11
US-10-213-990-70
; Sequence 70, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Storms, Reg
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
; FILE REFERENCE: 10182-019-999
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 1002
; TYPE: DNA
; ORGANISM: Aspergillus
US-10-213-990-70

Query Match 16.6%; Score 110; DB 14; Length 1002;
Best Local Similarity 57.8%; Pred. No. 1.8e-21;
Matches 216; Conservative 0; Mismatches 155; Indels 3; Gaps 1;

QY 201 CATAAATACGAGGCAACTTCCAAACCAAAATGGTAATGGTATTTATGCGTCTATGGTTG 260
Db 336 CATTACCTTCTCGGAGGTTCAATCCTTCGGGAAATGTTACCTGTCGGTGTATGGATG 395
QY 261 GACTGTGACCCCTTGTGGAATATTATTTGTCAGACAGTTGGGCA---ACTGGCGTCC 317
Db 396 GACTACCAACCCCTTAGTCGAATACTACATCTCGAGAATAATGCGAGTTACAACTCTGG 455
QY 318 ACCAGAGCAAGCCCTTAAGGAGCACTCACTGTTGATGAGGAAATATGATATCTAGA 377
Db 456 CTCGGGCATGACGCAACAGGGCACCGCTCACAGCATGATCCACTAGCATCTATGA 515
QY 378 GACTCTTAGAGTCAATCAACCCCTCAATTAAAGGGATGCGCATTTAAACAATATTGGAG 437
Db 516 GCACCAACAGGTCAACGAGCTTCGATCGTCGACGCGCCACTTCAACCAATATGTC 575
QY 438 TGTTTGAAGATCGAAACGCAAGTGGACAGATTTCTGTGACGACCACTTTAGAGCGTG 497
Db 576 CATCGGCCAAAACAAGCGATCCAGCGGCACAGTCAACCGCGAATCACTTCAAGGCCTG 635
QY 498 GGAACACTTAGGATCAATATCGGGAATATGATGAATCGCGCTTACTAGAGAGGCTA 557
Db 636 GCGTAGTCTGGGGATGAACCTGGGTACCCATAACTATCATGATTTTCCACTGAGGGATA 695
QY 558 TCAAGTAGCGGAA 571
```

```
Db 696 TGAGAGCAGCGGTA 709

RESULT 12
US-10-213-990-68
; Sequence 68, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Storms, Reg
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
; FILE REFERENCE: 10182-019-999
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Aspergillus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(705)
US-10-213-990-68

Query Match 16.4%; Score 108.8; DB 14; Length 705;
Best Local Similarity 51.9%; Pred. No. 3.5e-21;
Matches 298; Conservative 0; Mismatches 267; Indels 9; Gaps 2;

QY 6 CGTACCAGCAATTCCTATTGGCAACCAAGTGGCTAGTATGATTAATGAAATTTGGAAAGATAG 65
Db 120 CGGCACCCCAAGCTCCACCGGCTGGAAACAACGGCTACTACTCTCTTCTGGACTGATGG 179
QY 66 CGGTGGCTCTCGGCAATGATTCTCAATCATGCGGCTACGTTTCAGTGGCCCAATGGAACAA 125
Db 180 CGGCGGACGCTGACCTACCAATGGCGCGGTGGCTCTGCTACTCCGTCACCTGAGGANA 239
QY 126 TGTAAACAACATATTTATTCGTAAGGTAAATAATTCATGAAACACACAAACACACAAACA 185
Db 240 CGTGGGCAACTTTGTGGTGGAAAGGGC-----TGGAAACCTCGAAGCGCTAGGTACCG 293
QY 186 AGTTGTTAACTGTCATATAAAGTACGAGGCAACTTCCAAACCAATGTTATGCTATTT 245
Db 294 AGCTTTAAGTAGAACCATCAACTAGGAGGAGCTTCAACCCCGAGCGCAATGGCTACTCT 353
QY 246 ATGGCTCTATGTTGGACTGTTGACCCCTTCTTGTGCAATATTATATTGTCGACAGTTGGGG 305
Db 354 GGCTGTCTACCGCTGGACCAACACCCCTTGATTTGATGACTACGTTGTTGATGCTATGG 413
QY 306 CA---ACTGGCGTCCACGAGGCAACCGCTTAAGGGGACCATCACTGTTGATGAGGAGAAC 362
Db 414 TACATACAACCCCGGACGCGGTACTCTTCAGGGGCACCTGTCAACACCGCAGGTGGCAC 473
QY 363 ATATGATATCTAGAGACTCTTAGAGTCAATCAACCCCTCCATTAGGGGATGCCACATT 422
Db 474 TTACAACATCTACAGCGCGCTTCGCTACAATGCTCCCTCCATCGAAGGCAACCAAGACCTT 533
QY 423 TAAACAATATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCGACGATTTCTGTGAGCAA 482
Db 534 CACCCAGTACTGTTCTGTGGGCACTTCAAGGCTACCGGCGGCACTGTCCACCATGCCCAA 593
QY 483 CCACCTTTAGAGCGTGGGAAACTTAGGATGAATATGCGGAAATATGATGAGTCCGCT 542
Db 594 CCACCTTCAACGCTGGAGCAGACTGGGCATGAACCTCGGAACTCAACAACCTACAGATTGT 653
QY 543 TACTGTAGAGGCTATCAAGTAGCGGAGTGCT 576
Db 654 CGCCACTGAGGGTTACCAGAGCGGATCTGCT 687
```



## RESULT 13

US-10-299-393-1  
; Sequence 1, Application US/10299393  
; Publication No. US20030108642A1  
; GENERAL INFORMATION:  
; APPLICANT: Sabatier, Alain  
; APPLICANT: Fish, Neville Marshall  
; APPLICANT: Haigh, Nigel Paterson  
; TITLE OF INVENTION: PENICILLIUM FUNICULOSUM STRAIN USEFUL  
; TITLE OF INVENTION: FOR THE PRODUCTION OF ENZYMES  
; FILE REFERENCE: A32917-PCT-USA-1 (072667.0183)  
; CURRENT APPLICATION NUMBER: US/10/299,393  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: 09/462,246  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: EPO 9801161.5  
; PRIOR FILING DATE: 1998-05-06  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2898  
; TYPE: DNA  
; ORGANISM: Penicillium funiculosum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1317)...(1589)  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (1590)...(1642)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1644)...(2042)  
; FEATURE:  
; NAME/KEY: RBS  
; LOCATION: (570)...(576)  
; FEATURE:  
; NAME/KEY: RBS  
; LOCATION: (724)...(730)  
US-10-299-393-1

Query Match 16.0%; Score 106; DB 15; Length 2898;

Best Local Similarity 56.5%; Pred. No. 4.5e-20;

Matches 218; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

QY	195	CATGTCATATAA	CTACGGAGCCAACTTCCAAACCAATGGTAATGCGTATTATGCGGTCTA	254
DB	1640	CAGGACTGTCA	CGTACTCAGGAGAAATTTAACCCCTCTGGAAACGCTTATTGCGTGTCTA	1699
QY	255	TGFTTGACTGT	GACCCCTCTGTGCGAATATATATTTGTCGACAGTTGGGGCAACTGGCG	314
DB	1700	CGGGTGACAA	CAGATCCTCTGTGCGAATATATATCTCTGGAGTCTTACGGCACCTATAA	1759
QY	315	TCCACCA---	GGACCAACGCTTAAGGGGACCATCCTGTTGATGAGGAAACATATGATAT	371
DB	1760	CCCATCATCT	GACTTACTTCACTTGGCCAGGTCACTAGCGATGGTGGCACCTACGATAT	1819
QY	372	CTACGAGACT	CTTAGAGTCAATCAACCCCTCCATTAAGGGGATGCGCACATTTAAACAATA	431
DB	1820	CTACTCAACC	AGCGTGTCAACACAGCTTCCATTTAGGGAACCTTCCACCTTCAACACAGTA	1879
QY	432	TTGAGTGTTC	GAGATCGAAGAACGACGAGTGGCAGCATTTCTGTAGCAACCACTTTAG	491
DB	1880	CTGGTCACT	TCGACCCGAGAGCGAGTCGGCGGAACTGTCAACACGGCCAACTTTGC	1939
QY	492	AGCTTGGAAA	CACTTAGGATGATATGCGGAAATGTATCAAGTCGCGCTTACTGTAGA	551
DB	1940	AGCATGGA	AGGCACTTGGACTTGAATGGGCACTTATTAATGATGATTGTGTCACCGA	1999
QY	552	AGGCTATCAA	AGTAGTACGCGAAGTGTCTA	577
DB	2000	AGGCTACG	AGAGCAGTGGCTCTAGTA	2025

## RESULT 14

US-10-213-990-64  
; Sequence 64, Application US/10213990  
; Publication No. US20030082595A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Bussey, Howard  
; APPLICANT: Storms, Reg  
; APPLICANT: Roemer, Terry  
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL  
; TITLE OF INVENTION: ENZYMES AND METHODS OF USE  
; FILE REFERENCE: 10182-019-999  
; CURRENT APPLICATION NUMBER: US/10/213,990  
; CURRENT FILING DATE: 2002-08-05  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 64  
; LENGTH: 712  
; TYPE: DNA  
; ORGANISM: Aspergillus  
US-10-213-990-64

Query Match 15.8%; Score 104.8; DB 14; Length 712;

Best Local Similarity 55.8%; Pred. No. 5.2e-20;

Matches 221; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

QY	189	TGTTAATGTC	CCATAAATACGGAGCCAACTTCCAAACCAATGGTAATGCGTATTATG	248
DB	304	TGGATTCAG	AGCGGTCACTACAGCGGCTCTGGGAGACCGGNAACGGCTACTCTC	363
QY	249	CGTCTATG	TGGACTGTTGACCCCTTGTTCGAATATATATGTCGACAGTTGGGGCAA	308
DB	364	CGTGTACG	CTGACGACGAGTCCGCTGTCGAATCTTACATCGTGGAGAGTTACGGCTC	423
QY	309	CTGGCGTCC	ACCA---GGAGCAACGCTTAAGGGGACCATCTGTTGATGGAGGACATA	365
DB	424	CTATGACCC	CTCCACGGGAGCCACCATCTCGGCACCGTCGAGAGCGGCGGCCACGTA	483
QY	366	TGATATCT	ACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATGCGCACATTTAA	425
DB	484	CAACCTCT	ACAAGACGACCGGACGAAATGCGGCGTCCATCCAGGGACCGCTACTTTGA	543
QY	426	ACAATAAT	TCGAGTGTTCGAAGATCGAAACGACGAGTGGCAGCATTTCTGTAGCAACCA	485
DB	544	CCAGTACT	GGTTCGGTTCGGACTTCGCACCGGCAGAGTGGAACTGTGACGACGAAACCA	603
QY	486	CTTTAGAG	CGTGGAAAACTTAGGGATGAATATGGGAAAAATGTATGAAGTCGGCTTAC	545
DB	604	CTTTGATG	CGTGAGAAATGCGGGTCTGCAATTTGGGAACTTTTGAATATGATTGTTGC	663
QY	546	TGTAGAAG	CTATCAAGTAGCGGAAGTCTAATGT	581
DB	664	GACGGAGG	GGGTACCGAGCAGCGGCTCTGCTACTAT	699

## RESULT 15

US-10-213-990-65  
; Sequence 65, Application US/10213990  
; Publication No. US20030082595A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Bussey, Howard  
; APPLICANT: Storms, Reg  
; APPLICANT: Roemer, Terry  
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL  
; TITLE OF INVENTION: ENZYMES AND METHODS OF USE  
; FILE REFERENCE: 10182-019-999  
; CURRENT APPLICATION NUMBER: US/10/213,990  
; CURRENT FILING DATE: 2002-08-05  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 65

```

; LENGTH: 666
; TYPE: DNA
; ORGANISM: Aspergillus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(666)
US-10-213-990-65

```

```

Query Match      15.6%; Score 103.6; DB 14; Length 666;
Best Local Similarity 56.3%; Pred. No. 1.1e-19;
Matches 215; Conservative 0; Mismatches 164; Indels 3; Gaps 1;

Qy 203 TAAACTACGGAGCCAACTTCCAAACAAATGGTAATGCGTATTATTCGCTCTATGTTGGA 262
Db      |||||
Qy 272 TCACCTACAGCGGCTCCTGGCAGACACGCGGAACGGCTACCTCTCCGTGTACGGCTGGA 331
Db      |||||
Qy 263 CTGTTGACCCCTTTGTCGAATATTATATTCGACAGTTGGGGCAACTGGCGTCCACCA- 321
Db      |||||
Qy 332 CGACCAAGTCCGCTGGTTCGAATTTCTACATCGTGGAGAGTTACGGCTCTATGACCCCTCCA 391
Db      |||||
Qy 322 --GGAGCAACGCTAAGGGAGCACATCACTGTTGATGGAGGAACATATGATATCTACGAGA 379
Db      |||||
Qy 392 CGGGAGCCACCCATCTCGGCACCGTCGAGAGCGGGGCCACGTCACAACTCTACAAGA 451
Db      |||||
Qy 380 CTCTTAGAGTCAATCAACCCCTCCATTAAAGGGGATTGCCACATTTAAACAATTTGGAGTG 439
Db      |||||
Qy 452 CGACCGGAGCAATGCGCGCTCCATCCAGGGCAGCGCTACTTTGACCAGTACTGGTCGG 511
Db      |||||
Qy 440 TTCGAAGATCGAAACGACAGTGGCAGATTTCTGTGACGAACCACTTTAGAGCGTGGG 499
Db      |||||
Qy 512 TTCGGACTTCGACCGGCAGAGTGGAACTGTGACGACGAAGAACCACTTTGATGCGTGA 571
Db      |||||
Qy 500 AAACTTAGGGATGAATATGGGGAATCTATGAAGTCGCGCTTACTGTAGAGGCTATC 559
Db      |||||
Qy 572 GAAATGCGGGTCTGCAATTTGGGAACCTTTGACTATATGATTGTCGACGGAGGGGTACC 631
Db      |||||
Qy 560 AAAAGTAGCGGAAGTGTAAATGT 581
Db      |||||
Qy 632 AGAGCAGCGGCTCTGCTACTAT 653

```

Search completed: November 11, 2004, 02:08:22  
Job time : 425.895 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 16:24:32 ; Search time 2874.04 Seconds  
(without alignments)  
8406.125 Million cell updates/sec

Title: US-09-909-207-1  
Perfect score: 663  
Sequence: 1 CAATTCGTCACCGACAATTC.....TAATTTGGATAAAACAAT 663

Scoring table: IDENTITY\_NUC  
.Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98.6	14.9	618	7	CF472462 RTDS1_9 C
2	93	14.0	728	7	CN132941 OX1_9_D10
3	93	14.0	772	7	CN133022 OX1_9_D10
4	91.8	13.8	744	7	CF867983 trico13xe
5	91.8	13.8	799	6	CB898036 trico13xe
6	86.6	13.1	738	6	CD464145 ETH1_48 B
7	83.4	12.6	921	6	CD458837 Fg08_04B1
8	82.8	12.5	603	8	AQ160513 mgxb0006C
9	82.2	12.4	720	8	AQ447125 mgxb0005C
10	82.2	12.4	670	8	AQ361561 mgxb0004B
11	82.2	12.4	750	8	AQ160254 mgxb0003L
12	81.6	12.3	786	8	AQ325248 mgxb0021M
13	81.4	12.3	583	8	AQ399120 mgxb0001B
14	76.6	11.6	561	8	AQ396475 mgxb00010M
15	76.4	11.5	617	1	AJ638869 AJ638869
16	76.2	11.5	768	7	CF881056 trico83xj
17	76.2	11.5	822	6	CB907827 trico83xj
18	74.6	11.3	753	8	AQ448084 mgxb0016B
19	74.4	11.2	418	8	AQ398756 mgxb0005L
20	71.8	10.8	746	6	CB901964 trico28xi
21	71.8	10.8	746	7	CF871731 trico28xi
22	69.8	10.5	520	5	BQ165950 WHE0821-0
23	66.2	10.0	561	5	BQ471960 HV04A02r
24	63.2	9.5	473	6	CD464005 ETH1_48_B

C	25	60.4	9.1	483	5	BQ664593
C	26	53.6	8.1	646	8	AQ449078
C	27	51.6	7.8	493	7	CO135442
C	28	49.4	7.5	490	2	BF200865
C	29	49.4	7.5	532	4	BM134812
C	30	48.4	7.3	532	7	CN008236
C	31	47.8	7.2	743	7	CF865758
C	32	47.8	7.2	802	6	CB895680
C	33	46.6	7.0	503	5	BQ166480
C	34	46.6	7.0	1582	9	CC820765
C	35	46.4	7.0	456	1	AJ637947
C	36	45.2	6.8	587	6	CD033274
C	37	44.8	6.8	614	6	CA278095
C	38	41.6	6.3	488	4	BM135798
C	39	40.4	6.1	386	7	CN470894
C	40	40.2	6.1	665	2	BB664921
C	41	40.2	6.1	1319	3	AK053390
C	42	39.8	6.0	363	6	CD113872
C	43	39.6	6.0	704	5	BQ110128
C	44	38.8	5.9	588	8	AQ630020
C	45	38.8	5.9	750	6	CA447770

## ALIGNMENTS

RESULT 1  
CF472462  
LOCUS  
DEFINITION  
RTDS1\_9\_C03.g1\_A015 Drought-stressed loblolly pine roots DS1 Pinus taeda cDNA clone RTDS1\_9\_C03\_A015 5', mRNA sequence.  
CF472462  
CF472462.1 GI:34489834  
EST.  
Pinus taeda (loblolly pine)  
Pinus taeda  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 618)  
Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.M., Dean, J., Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and Neale, D.  
An EST database from drought-stressed loblolly pine (Pinus taeda) roots  
Unpublished (2003)  
Other ESTs: RTDS1\_9\_C03.b1\_A015  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@cuga.edu  
RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
Seq primer: JENREV (CAGGAACAGCTATGACC).

FEATURES  
source  
Location/Qualifiers  
1..618  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="CCLONES"  
/db\_xref="taxon:3352"  
/clone="RTDS1\_9\_C03\_A015"  
/lab\_host="DH10B-T1-phage-resistant E. coli"  
/clone\_lib="Drought-stressed loblolly pine roots DS1"  
/notes="Vector: pSL180; Site\_1: EcoRI; Site\_2: XhoI; The library was prepared from polyA+ RNA from drought-stressed loblolly pine (Pinus taeda) roots. Water was withheld from

ramet clones until predawn needle water potential reached -1.75 MPa. On day 7 roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSul180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

## ORIGIN

```

Query Match      14.9%; Score 98.6; DB 7; Length 618;
Best Local Similarity 57.4%; Pred. No. 4.8e-17;
Matches 198; Conservative 0; Mismatches 144; Indels 3; Gaps 1;

QY 182 AACAGTTGGTAACATGTCCATAACTAGGAGCCAACTTCCAAACCAATGGTAATGGT 241
DB 254 ATCCAGGTAGTGGCAGACATCACTTCTCGGTACTTTCAGCCCAAGAGAAAGCT 313
QY 242 ATTTAAGCGTCTATGGTGTGAGTGTGACCTCTTGTGCAATATATATATGTGACAGTT 301
DB 314 ATCTTGGCGTCTACGGCTGGACCAAGCCCTCTGATGGAATACATCTACATTCTCGAAAGCT 373
QY 302 GGGCAACTGGCGTCC---ACCAGAGCAACGCCCTAAGGGACCATCACTCTTATGGAG 358
DB 374 ATGGGACCTCAACACCTCGATCCAGCATGACCCATATGGGCATGTCCAGGCGAGGTG 433
QY 359 GAACATATGATCTATACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAAGGGGATTGCCA 418
DB 434 CGACCTATGATATCTATGACCAAGCAAGTCAACAGCCCTCTATCTCGCGCAGGCGA 493
QY 419 CATTAAACATATTTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGCATTTCTGTCA 478
DB 494 CTTTCAACCAATCTATGTTCCATCCCTCAGACGACGCGTCCAGCGGAATGTCACTACAG 553
QY 479 GCAACCACTTTAGACGCTGGGAAACTTAGGATGAATATGGGA 523
DB 554 CGAACCAITTTCAACGCTGGGTGCTCTTGGGCATGAACCTGGGA 598

```

## RESULT 2

```

CN132941
LOCUS OX1_9_D10_b1_A002 Oxidatively-stressed leaves and roots Sorghum
DEFINITION bicolor cDNA clone OX1_9_D10_A002 3', mRNA sequence.
ACCESSION CN132941
VERSION CN132941.1 GI:45963408
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 728)
Cordonnier-Pratt, M.-M., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C.,
Sun, F., Sullivan, R., Lim, S., Eastman, A., and Pratt, L.H.
An EST database from Sorghum: oxidatively stressed leaves and roots
Unpublished (2003)
Other ESTs: OX1_9_D10_A002
Contact: Cordonnier-Pratt, M.M.
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq primer: Sug3-14 (TAGTCTAGCGCGCGGAC)
POLYA=Yes.
Location/Qualifiers

```

## FEATURES

## source

```

1..728
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_cultivar="Brix623"
/db_xref="taxon:4558"
/clones="OX1_9_D10_A002"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="Oxidatively-stressed leaves and roots"
/note="Organ: Leaf and Root; Vector: pME18S-FL3; Site: 1:
XhoI; Site 2: XhoI; The library was prepared from polyA+
RNA from oxidatively stressed, hydroponically grown
sorghum seedlings. At 8 days of age, growth medium was
supplemented with hydrogen peroxide to 0.003% and leaves
were misted with 10 uM methyl viologen. Leaves and roots
were harvested at 3, 12 and 27 hr after treatment and all
tissue pooled. Double-stranded cDNA was cloned
unidirectionally into different DraIII sites of the
pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG,
3-prime DraIII site is CACCATGTG). XhoI excises the cDNA
insert."

```

## ORIGIN

```

Query Match      14.0%; Score 93; DB 7; Length 728;
Best Local Similarity 53.9%; Pred. No. 2e-15;
Matches 214; Conservative 0; Mismatches 180; Indels 3; Gaps 1;

QY 200 CCATAAACTAGGAGCCAACTTCCAAACCAATGGTAATGGTCTATGGTCTATGGTT 259
DB 101 CCATCAACTATGGCGGTTCTTTAGCCCTCAGGTAACGGCTACTCTGCGTCTACGGCT 160
QY 260 GGACTGTTGACCCCTCTTGTGCAATATATATTTGCGACAGTTGGGGCA---ACTGGCGTC 316
DB 161 GGACTCGCAGCCCTCTGTTGAGTACTACGTACGAGAACTACGGCACTTACAACCCGTG 220
QY 317 CACCAGGAGCAACGCCCTTAAGGGACCATCACTGTTGATGGAGGAACATATGATATCTACG 376
DB 221 GCTCTGTGGCCAGCAACAAAGGGCACCGCTCTACAACGAGCGGACACCTACCATCTTACC 280
QY 377 AGACTCTTAGAGTCAATCAACCCCTCCATTAAAGGGGATTGCCACATTTAAACAATTTGGA 436
DB 281 AGACACCCGCTACAAACAGCCCTCTTATCGAGGCCACAGACCTTCAACAGTACTGGG 340
QY 437 GTGTTTGAAGATCGAAACGACGAGTGGCAGATTTCTGTGCAACAACCATTTAGAGCGT 496
DB 341 CCATCGCGCCCAACAAGCGCAGCGCGCGTCAACATCGACACTATCTTCAATGCTT 400
QY 497 GGGAAAATTTAGGATGAATATGGGAAATGTATGAAGTCGCGTCTACTAGAGGCT 556
DB 401 GGGCTAAGCGTGGTATGAGACTTGGAAACCACTACTACAGATCTTGCTGCTACCGAGGAT 460
QY 557 ATCAAGTAGCGGAAGTGCTAATGTATATAGCAATAC 593
DB 461 ACCAGCAGTGGATCTTCTTCTATCTATGTCCAGAC 497

```

## RESULT 3

```

CN133022
LOCUS OX1_9_D10_g1_A002 Oxidatively-stressed leaves and roots Sorghum
DEFINITION bicolor cDNA clone OX1_9_D10_A002 5', mRNA sequence.
ACCESSION CN133022
VERSION CN133022.1 GI:45963542
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 772)
Cordonnier-Pratt, M.-M., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C.,
Sun, F., Sullivan, R., Lim, S., Eastman, A., and Pratt, L.H.
An EST database from Sorghum: oxidatively stressed leaves and roots
Unpublished (2003)

```

## COMMENT

Other ESTs: OX1.9.D10.b1.A002  
 Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmprat@uga.edu  
 Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in  
 the Human Genome Center, University of Tokyo Institute of Medical  
 Science; plant material and RNA prepared at Texas A & M University;  
 sequencing done in the Laboratory for Genomics and Bioinformatics,  
 University of Georgia. Sequence ends have been trimmed to exclude  
 vector and regions below Phred quality 16. Three-prime sequences  
 are presented as their reverse complement and have been trimmed to  
 exclude polyA.  
 Seq primer: Sug5 (CTTGTGCTTAAAGCTGGC).

## FEATURES

source

1. .772  
 Location/Qualifiers  
 /organism="Sorghum bicolor"  
 /mol\_type="mRNA"  
 /cultivar="BTx623"  
 /db\_xref="taxon:4558"  
 /clone="OX1.9.D10.A002"  
 /lab\_host="DH10B-Ti phage-resistant E. coli"  
 /clone\_lib="Oxidatively-stressed leaves and roots"  
 /note="Organ: Leaf and Root; Vector: pME18S-FL3; Site 1:  
 XhoI; Site 2: XhoI; The library was prepared from polyA+  
 RNA from oxidatively stressed, hydroponically grown  
 sorghum seedlings. At 8 days of age, growth medium was  
 supplemented with hydrogen peroxide to 0.003% and leaves  
 were misted with 10 uM methyl viologen. Leaves and roots  
 were harvested at 3, 12 and 27 hr after treatment and all  
 tissue pooled. Double-stranded cDNA was cloned  
 unidirectionally into different DraIII sites of the  
 pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG,  
 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA  
 insert."

## ORIGIN

Query Match 14.0%; Score 93; DB 7; Length 772;  
 Best Local Similarity 53.9%; Pred. No. 2e-15;  
 Matches 214; Conservative 0; Mismatches 180; Indels 3; Gaps 1;  
 QY 200 CCATAAAGTACGAGGCACTTCCAAACCAATGGTATGCGTATTTATGCGTCTATGGTT 259  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 324 CCATCAACTATGCGGGTCTTTCAGCCCTCAGGGTAAACGGCTACCTCTGCGTCTACGGCT 383  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 260 GGAAGTGTGACCTCTGTGCAATATTTATTTGCGACAGTTGGGCA---ACTGGCGTC 316  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 384 GGAAGTGTGACCTCTGTGCAATATTTATTTGCGACAGTTGGGCA---ACTGGCGTC 443  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 317 CACGAGGAGCAAGCCCTTAAAGGGACCATCTGTGTGATGGAGGAACATATGATCTACG 376  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 444 GCTCTGCTGCCAGGACAGGGCACCCTGTACAAGCAGCGGACACCTACGATCTCTACC 503  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 377 AGACTCTTAGAGTCAATCAACCTCTCAATTAAGGGATGGCCATTAAACATATTTGA 436  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 504 AGACACCCCGCTACACACGCGCTCTATCGACGGCCCAACAGACCTTCAACGAGTACTGGG 563  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 437 GTGTTGGAAGATCGAAGCAGCAGGTGGCAGGATTTCTGTGAGCAACCTTTAGAGCGT 496  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 564 CCATCCCGCGCAACAGCGGACGCGCGCGCTCAACATGCGACATCTTCAATGCTT 623  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 497 GGGAAACCTTAGGATCAATATGGGAAATGTATGAAGTTCGCGCTTACTGTAGAAGGCT 556  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 624 GGGCTAACGCTGGTATGAGACTTGGAAACCACTACTACCAGATCTCTGGCTACCGAGGAT 683  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 557 ATCAAAGTAGCGGAAGTCTTAATGTATATAGCAATAC 593  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 684 ACCAGACAGTGGATCTTCTTATCTATCTATGTGCCAGAC 720

## RESULT 4

CF867983  
 LOCUS  
 DEFINITION  
 CF867983  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 FEATURES  
 source  
 ORIGIN  
 Query Match 13.8%; Score 91.8; DB 7; Length 744;  
 Best Local Similarity 55.0%; Pred. No. 4.5e-15;  
 Matches 202; Conservative 0; Mismatches 162; Indels 3; Gaps 1;  
 QY 193 AACATGTCATAAACTACGAGCCAACTTCCAAACCAATGGTATGCGTATTTATGGCTC 252  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 353 AACAAAGGTCAATCAACTTCTCGGCGAGCTACAAACCCCAACGGCAACAGCTACCTCTCGGTG 412  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 253 TATGGTTGAGTGTTCACCTCTTGTGGAATATTTATTCGACAGTTGGGCGCAACTGG 312  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 413 TAGCGGTGTCGCGCAACCCCTTGATGAGTACTACATGTCGAGAACTTTGGCACCTAC 472  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 313 CGTCCACCA---GGAGCAACGCTTAAGGGAGCACTCACTGTTGATGGAGGAACATATGAT 369  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 473 AACCGGTCCACGGCGCCACCAAGCTGGGCGAGGTCACTCCGACGCGCTACGAC 532  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 370 ATCTACGAGACTCTTAGAGTCAATCAACCTCTCAATTAAGGGGATTTGCCACATTTAAACAA 429  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 533 ATTTACCGCAGCAGCGCGCTCAACACCGCTCCATCATCGGCACCGCCACTTTTACAG 592  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 430 TATTGAGTGTTCGAAGATCGAAACGACGAGTGGCAGGATTTCTGTCAACACCACTTTT 489  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 593 TACTGTGCTGTCGCGCGCAACCCCGCTCGAGCGGCTCCGTCACACGGGAGAACCTTC 652  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 490 AGAGCGTGGGAAACTTAGGGATGAATATGGGGAATAATGTATGAAGTCGCGCTTACTGTGA 549  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 653 AACCGGTGGGCTCAGCAAGGCTGACGCTCGGAGCGATGATACCAGATTGTTGCCGTG 712  
 QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 550 GAAGGCT 556

CF867983 744 bp mRNA linear EST 31-OCT-2003  
 tricol3xe09.bl T.reesei mycelial culture, Version 6 October 2003  
 Hypocrea jecorina cDNA clone tricol3xe09, mRNA sequence.

CF867983.1 GI:38122635

EST.  
 Hypocrea jecorina (anamorph: Trichoderma reesei)  
 Hypocrea jecorina  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.

1 (bases 1 to 744)  
 Diener S.E., Dunn-Coleman N., Foreman P., Houfek T.D.,  
 Teunissen P.J.M., van Solingen P., Dankmeyer L., Mitchell T.K.,  
 Ward M. and Dean R.A.

Characterization of the protein processing and secretion pathways  
 in a comprehensive set of expressed sequence tags from Trichoderma  
 reesei

FEMS Microbiol. Lett. 230 (2), 275-282 (2004)

Contact: Ralph A. Dean  
 Fungal Genomics Laboratory  
 North Carolina State University  
 Campus Box 7251, Raleigh, NC 27695, USA

Tel: 919-513-0020

Fax: 919-513-0024

Email: ralph.dean@ncsu.edu

Seq primer: Tr-P1 primer.

Location/Qualifiers

1. .744

/organism="Hypocrea jecorina"

/mol\_type="mRNA"

/strain="QM6a"

/db\_xref="taxon:51453"

/clone="tricol3xe09"

/dev\_stage="mycelia"

/clone\_lib="T.reesei mycelial culture, Version 6 October 2003"

/note="Vector: pREP3Y; Site 1: Not I/Sal I; Mycelial

culture grown from 24 hrs to 6 days with varying Carbon

and Nitrogen sources and concentrations."

```

Db      713 GGAGGCT 719

RESULT 5
CB898036      799 bp      mRNA      linear      EST 02-JUL-2003
LOCUS      tricol3xe09 T.reesei mycelial culture, Version 3 april Hypocrea
DEFINITION      jecorina cDNA clone tricol3xe09, mRNA sequence.
ACCESSION      CB898036
VERSION      CB898036.1 GI:30112694
KEYWORDS      EST.
SOURCE      Hypocrea jecorina (anamorph: Trichoderma reesei)
ORGANISM      Eukaryota; Fungi; Ascomycota; Peizomycotina; Sordariomycetes;
REFERENCE      Hypocreales; Hypocreales; Hypocreaceae; Hypocrea.
AUTHORS      1. (bases 1 to 799)
Foreman,P.K., Brown,D.E., Dankmeyer,L., Dean,R., Diener,S.,
Dunn-Coleman,N.S., Goedegebuur,F., Houfek,T.D., England,G.J.,
Kelley,A.S., Meerman,H.J., Mitchell,T., Mitchinson,C.,
Olivares,H.A., Teunissen,P.J., Yao,J., and Ward,M.
TITLE      Transcriptional regulation of biomass-degrading enzymes in the
JOURNAL      filamentous fungus Trichoderma reesei
MEDLINE      J. Biol. Chem. 278 (34), 31988-31997 (2003)
PUBMED      22803314
COMMENT      12788920
Contact: Pamela K. Foreman
Genencor Intl.
925 Page Mill Road, Palo Alto, CA 94304, USA
Tel: (650) 846-7635
Fax: (650) 621-7817
Email: Pforeman@genencor.com
Seq primer: LT-F1 primer.
FEATURES
Location/Qualifiers
1..799
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/clone="tricol3xe09"
/dev_stages="mycelia"
/clone_lib="T.reesei mycelial culture, Version 3 april"
/note="Vector: pREP3y; Site 1: Not I/Sal I; Mycelial
culture grown from 24 hrs to 6 days with varying Carbon
and Nitrogen sources and concentrations."

ORIGIN
Query Match      13.8%; Score 91.8; DB 6; Length 799;
Best Local Similarity 55.0%; Pred. No. 4.5e-15;
Matches 202; Conservative 0; Mismatches 162; Indels 3; Gaps 1;

QY      193 AACATGTCATTAAGTACGAGGACCACTTCCAAACCAATGGTAATGCGTATTATGCGTC 252
Db      408 AACAAAGGTATCAACTCTTCGGGCGAGCTACACCCCAACGGCAAGTACTCTCCGTG 467
QY      253 TATGTTGGACGTGTTGACCCCTCTTGCGAATATTATTGTCGACAGTTGGGGCAACTGG 312
Db      468 TACGGCTGTCCCGCAACCCCTGATCGAGTACTACATGTCGAGAACTTTGGCACTTAC 527
QY      313 CGTCCACCA--GGAGCAAGCGCTAAGGGGACATCATCTGTTGATGGAGGAACATATGAT 369
Db      528 AACCCGCTCCACGGGGCCCAACGAGTGGGCGAGGTCACTCCGACGGGCGGTCTACGAC 587
QY      370 ATCTAGAGACTTTAGAGTCAATCAACCTCATTAAGGGGATGTCACATTTAAACAA 429
Db      588 ATTATCCGACGAGCGCGCTCAACAGCGCTCAATCATCGGACCGCCACTTTTACGAG 647
QY      430 TATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGATTCTTGTGAGCAACCACTTT 489
Db      648 TACTGTGCTCGCGCGCAACCAACCGCTCGAGCGGTCCGTCAACACGCGCAACCACTTC 707
QY      490 AGAGCGTGGAAACCTTAGGATGAATATGCGGAAATGTATCAAGTCGCGCTTACTGTA 549
Db      708 AACCGTGGGCTCAGCAAGGCGCTGACGCTCGGACGATGGATTACCAAGATTGTGCCG 767

```

---

```

QY      550 GGAGGCT 556
Db      768 GGAGGCT 774

RESULT 6
CD464145      738 bp      mRNA      linear      EST 04-JUN-2003
LOCUS      ETH1_48_B06.g1 A002 Ethylene-treated seedlings Sorghum bicolor cDNA
DEFINITION      clone ETH1_48_B06_A002 5', mRNA sequence.
ACCESSION      CD464145
VERSION      CD464145.1 GI:31385413
KEYWORDS      EST.
SOURCE      Sorghum bicolor (sorghum)
ORGANISM      Sorghum bicolor.
REFERENCE      Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 738)
Cordonnier-Pratt,M.-M., Wentzel,V., Suzuki,Y., Sugano,S.,
Klein,R.R., Liang,C., Sun,F., Sullivan,R., Shah,M., Salzman,R.,
Chua Tan,N., Gonzalez,M., Lane,S., Miller,V., Nanda,P.,
Olaseinde,O., Eastman,A. and Pratt,L.H.
TITLE      An EST database from Sorghum: 1-aminocyclopropane-1-carboxylic acid
JOURNAL      (ACC)-treated seedlings
COMMENT      Other_ESTs: ETH1_48_B06.bl_A002
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq primer: Sug5 (CTTCTGCTCTAAAGCTCGC).
FEATURES
Location/Qualifiers
1..738
/organism="Sorghum bicolor"
/mol_type="mRNA"
/cultivar="BTx623"
/db_xref="taxon:4558"
/clone="ETH1_48_B06_A002"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="Ethylene-treated seedlings"
/note="Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The
library was prepared from polyA+ RNA from seedlings grown
in hydroponic culture. At 8 days of age, medium was
supplemented with 0.5 mM 1-aminocyclopropane-1-carboxylic
acid (ACC) to induce endogenous ethylene (ETH) production.
Roots and shoots were harvested after 27 and 72 hr and
material from both time points was combined prior to RNA
isolation. Double-stranded cDNA was cloned
unidirectionally into different DraIII sites of the
pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG,
3-prime DraIII site is CACCATGTG). XhoI excises the cDNA
insert."

ORIGIN
Query Match      13.1%; Score 86.6; DB 6; Length 738;
Best Local Similarity 53.5%; Pred. No. 1.4e-13;
Matches 204; Conservative 0; Mismatches 174; Indels 3; Gaps 1;

QY      200 CCATTAACCTACGAGCCCACTTCCAAACCAATGGTAATGCGTATTATGCGTCTATGTT 259

```

Db 345 CCATCAAGTACTCTGCTACTTACAACATCAACGAAACAGCTACCTCGCTGTTTACGGAT 404  
 QY 260 GGACTGTTGACCCCTTTGTCGAATATTATTTGTCGACAGTTGGGGCACTGGCGTCCAC 319  
 Db 405 GGACTCAGAACCCCTCTCATCGAGTACTACATCGTTTGAAGTTCGGCACTTCAACCCCT 464  
 QY 320 CA--GGACCAACCGCTAAGGGGACCACATCACTGTTTGTATGGAGAACATATATCTACG 376  
 Db 465 CTTCCGGCGCGCAGAGAAGGTGAGGTCACTGTTGACGGATCTGTCTACGACATCTACG 524  
 QY 377 AGACTCTTAGAGTCATCAACCCCTCCATTAAGGGGATTCGCCATTTAAACAATATTGA 436  
 Db 525 TCAGCAGCCCGTGTCAACGCCCTCCATTTAGGGGTAAACAAGACCTTTTACGACAGTTCTGT 584  
 QY 437 GTGTTGGAAGATCGAAACGACGAGTGGCAGCATTTCTGTCTAGCAACCACTTTAGAGCGT 496  
 Db 585 CTGTTTCGACGCAACAGCGATCCAGCGGATCCGTCACACCGGTGCTCACTTCCAGGCT 644  
 QY 497 GGGAAACTTAGGATGAATATGGGGAATATGTAAGTCCGCGCTTACTGTAGAGGCT 556  
 Db 645 GGAATAATGTGGCTCACTTGGTAAACCAACTATCAGATCTTGTCTGTGAGGCT 704  
 QY 557 ATCAAGTAGCGGAAGTGCTA 577  
 Db 705 ACTACAGCTCGGCTCTGCCA 725

RESULT 7  
 CD458837 921 bp mRNA linear EST 14-JUN-2004  
 LOCUS Fg08\_04b10\_A Fg08\_AAPC\_ECORC\_Fusarium\_graminearum\_complex\_substrate  
 DEFINITION Gibberella zeae cDNA clone Fg08\_04b10, mRNA sequence.  
 CD458837  
 EST.  
 CD458837.1 GI:31373577

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Gibberella zeae  
 Gibberella zeae  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.  
 1 (bases 1 to 921)  
 Watson, R.J., Heyes, R., Chapados, J., Couroux, P., Harris, L.J.,  
 Hattori, J., Lacroix, C., Ouellet, T., Robert, L.S., Singh, J.A.,  
 Spratt, D. and Tinker, N.A.  
 A cDNA library prepared from *Fusarium graminearum* grown on a  
 complex plant substrate  
 Unpublished (2003)  
 Contact: Watson, Robert.J.  
 Eastern Cereal and Oilseed Research Centre  
 Agriculture and Agri-food Canada  
 Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,  
 CANADA  
 Tel: (613) 759-1655  
 Fax: (613) 759-1701  
 Email: watsonrj@agr.gc.ca.

FEATURES  
 source  
 1..921  
 /organism="Gibberella zeae"  
 /mol\_type="mRNA"  
 /strain="DAOM 180378"  
 /db\_xref="taxon:5518"  
 /clone="Fg08\_04b10"  
 /tissue\_type="Mycelium"  
 /dev\_stage="Asexual"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="Fg08\_AAPC\_ECORC\_Fusarium\_graminearum\_complex\_s  
 ubstrate"  
 /note="Vector: pBluescript II+; Site 1: EcoRI; Site 2:  
 XhoI; *Fusarium graminearum* grown on a complex plant  
 substrate-- wheat leaves treated to remove most of the low  
 molecular weight, water-soluble components."

ORIGIN  
 Query Match 12.6%; Score 83.4; DB 6; Length 921;

Best Local Similarity 52.4%; Pred. No. 1.2e-12;  
 Matches 208; Conservative 0; Mismatches 186; Indels 3; Gaps 1;  
 QY 200 CCATTAATCTAGGACCAACTTCCAAACCAATGTAATGCGTATTTATCGCTCTATGGTT 259  
 Db 363 CCATCAATCTAGGAGGTTCCTTCAACCTTCAGGGTAACGGATACCTTTCCGTTTACGGAT 422  
 QY 260 GGACTGTTGACCCCTCTGTCGAATATTATTTGTCGACAGTTGGGGCACTGGCGTCCAC 319  
 Db 423 GGACCCGGTCCCTCGTCGAGTACTACGTATCGAGAGTTACGGTTCTTACAACCCCG 482  
 QY 320 CAGGA---GCAACGGCTAAGGGGACCATCACTGTTGATGGAGAACATATATCTACG 376  
 Db 483 CGACCGAGGCTCAGCACCGAGGTACCGTCTACACCGCGTGACACCTACGATCTCTATA 542  
 QY 377 AGACTCTTAGAGTCATCAACCCCTCCATTAAGGGGATTCGCCACATTTAAACAATATTGA 436  
 Db 543 TGTCCACCGCTTACCAACAGCCTTCATGACGGTGTTCAGACCTTCAACAGTACTGGT 602  
 QY 437 GTGTTGGAAGATCGAAACGACGAGTGGCAGCATTTCTGTCTAGCAACCACTTTAGAGCGT 496  
 Db 603 CCATCGCGCGCAACAGCGTACCAGCGCTCCGTCAACATGCAAGAACCACTTCAATGCTT 662  
 QY 497 GGGAAACTTAGGATGAATATGGGGAATATGTAAGTCCGCGCTTACTGTAGAGGCT 556  
 Db 663 GGAGATCTGCTGGCATGAACCTCGGAAACCACTACTACCAGATTCTGGCCACTGAGGTT 722  
 QY 557 ATCAAGTAGCGGAAGTGCTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 593  
 Db 723 ACCAGAGCAGTGGCTCATCTTCTATCTATGTCAGAC 759

RESULT 8  
 LOCUS AQ160513/c  
 DEFINITION mgxb0006C21r CUGI Rice Blast BAC library Magnaporthe grisea genomic  
 clone mgxb0006C21r, genomic survey sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Magnaporthe grisea (anamorph: Pyricularia grisea)  
 Magnaporthe grisea  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Sordariomycetes Incertae sedis; Magnaportheaceae; Magnaporthe.  
 1 (bases 1 to 603)  
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,  
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.  
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea  
 Genome  
 Unpublished (1998)  
 Contact: Dean RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson University, Clemson, SC 29634  
 Tel: 864 656 5737  
 Fax: 864 656 4293  
 Email: rdean@clemson.edu  
 Seq primer: GGAACAGCTATGACCATG  
 Class: BAC ends  
 High quality sequence stop: 424.  
 Location/Qualifiers

FEATURES  
 source  
 1..603  
 /organism="Magnaporthe grisea"  
 /mol\_type="genomic DNA"  
 /strain="70-15"  
 /db\_xref="taxon:148305"  
 /clone="mgxb0006C21r"  
 /tissue\_type="Protoplasts"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="CUGI Rice Blast BAC Library"  
 /note="Vector: pBACWICH; Site\_1: HindIII; Site\_2: HindIII;  
 Rice blast is one of the most devastating fungal diseases  
 of rice world wide. It is a filamentous ascomycete with

a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

ORIGIN

Query Match 12.5%; Score 82.8; DB 8; Length 603;  
 Best Local Similarity 52.8%; Pred. No. 1.6e-12;  
 Matches 227; Conservative 0; Mismatches 197; Indels 6; Gaps 2;  
 QY 154 AAAAAATTCAATGAACACAAACACCAACAAAGTTGGTAATCATGTCATTAATACGGA 213  
 Db 444 AAGAAAAAAGAGACTACCAACAAACAAACAAACAGCCGCTCATCAACTACTCG 385  
 QY 214 GCCAACTTCCACCAAGTAATGCGTATTTATGCGTCTATGTTGACTGTTGACCT 273  
 Db 384 GGCAGCTACAGCCCGGCGCAACTCATACCTGGCCGCTACGGCTGACCGGCAACCG 325  
 QY 274 CTGTGCAATATATATGTCACAGTTGGGGCACTGGCGTCCACCA---GGAGCAACG 330  
 Db 324 CTGATCGAGTACTAGTGTGGAGAGCTTTGGCAGCTACACCCGCTGTCGGCGCCACC 265  
 QY 331 CTTAAGGGAGCATCATCTGTTGATGGAGAACATATGATATCTACGAGCTCTTAGAGTC 390  
 Db 264 AACCCTGGGCTCTTCACTCTGGAGCGGCACTACGACATCTCTGGTCTGAGCACCCTG 205  
 QY 391 ATCAACCTCCATTAAGGGGATTCACATTTAAACAATATGAGTGTTCGAAGATCG 450  
 Db 204 AACGAGCCCTCATCGAGGCAACGACCTTTGCGAGCTCTGCTGCGTGGCGCGAAC 145  
 QY 451 AAACGACAGTGGGACGATTTCTGTCAGCAACACCTTTAGAGCGTGGGAAACTTAGGG 510  
 Db 144 AAGCGGCGGAGCGGACCGTCACTTTGCGCAACACGCTCAACGCTGCGCAACCGCGC 85  
 QY 511 ATGATATGGGAAATATGATGAAGTCGG---CTTACTGTAGAAGGCTATCAAGTAGC 567  
 Db 84 CTCAACCTCGGCAACAGTGGAACTACCAAGTCTCTGACCGTCTGAGGGCTACCAAGCAGC 25  
 QY 568 GGAAGTGCTA 577  
 Db 24 GCTCGGCCA 15

RESULT 9  
 AQ47125/c  
 LOCUS  
 DEFINITION  
 clone mgxb0005C20f, genomic survey sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Magnaporthe grisea (anamorph: Pyricularia grisea)

REFERENCE  
 AUTHORS  
 TITLE  
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome

JOURNAL  
 COMMENT  
 Unpublished (1998)  
 Contact: Dean RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson University, Clemson, SC 29634  
 Tel: 864 656 5737  
 Fax: 864 656 4293  
 Email: rdean@clemson.edu  
 Seq primer: TAATACGACTCTACTATAGGG

Class: BAC ends  
 High quality sequence stop: 263.  
 Location/Qualifiers  
 1. 670

/organism="Magnaporthe grisea"  
 /mol\_type="genomic DNA"  
 /strain="70-15"  
 /db\_xref="taxon:148305"  
 /clone="mgxb0005C20f"  
 /tissue\_type="Protoplasts"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="CUGI Rice Blast BAC Library"  
 /notes="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;  
 Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

ORIGIN

Query Match 12.4%; Score 82.2; DB 8; Length 670;  
 Best Local Similarity 54.6%; Pred. No. 2.5e-12;  
 Matches 209; Conservative 0; Mismatches 168; Indels 6; Gaps 2;  
 QY 201 CATAAATACGGAGCCAACTTCCAAACCAATGGTATGCGTATTTATGCTCTATGTTG 260  
 Db 397 CATCACTACTCGGGCAGCTACAGCCCCAGGGCACTCATCCTGGCCGCTCTACGGCTG 338  
 QY 261 GACTGTTGACCTCTTTCGAATATATTTATGTCAGAGTTGGGCACTGGGCTCCACC 320  
 Db 337 GACCGCAACCCGCTGATCGAGTACTAGTGTGGAGAGCTTTGGAGCTACACCGCTC 278  
 QY 321 A---GGAGCAACCGCTTAAGGGGACCATCCTGTTTGTAGGAGGAACATATGATATCTACGA 377  
 Db 277 GTCCGGCGCCACCAACCGGGGCTCTTCACTCGAGCGGCGAGCACTTACGACATCTCGT 218  
 QY 378 GACTCTTAGAGTCAATCAACCTCCATTAAGGGAATGCCACATTTAAACAATATTCGAG 437  
 Db 217 CAGCACCCGCTTACCAACCAAGCCCTCCATCGAGCGGCAACCAAGACCTTTTCAGCAGTCTCTGTC 158  
 QY 438 TGTTCGAAGATCGAAACGACGAGTGGCAGCATTTCTGTCAGCAACCACTTTAGAGCGTG 497  
 Db 157 GGTGCGCGCAACGAGCGGCGCAGCGCACGTCACCTTTGCCAACCAACGCTCAACGCTG 98  
 QY 498 GGAATACTTAGGATGAATATGGGAAATATGAAAGTCCG---CTTACTGTAGTAGG 554  
 Db 97 GCGCAACCGCGGCTCAACCTCGGCAACCAAGTGGAACTTACCAGATCTCTGCGCGTCGAGGG 38  
 QY 555 CTATCAAGTAGCGGAGTGCTA 577  
 Db 37 CTACCACGACGCGGCTCCGCCA 15

RESULT 10  
 AQ361561/c

LOCUS  
 DEFINITION  
 clone mgxb0004B19r, genomic survey sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Magnaporthe grisea (anamorph: Pyricularia grisea)

REFERENCE  
 AUTHORS  
 TITLE  
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome

JOURNAL  
 COMMENT  
 Unpublished (1998)  
 Contact: Dean RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson University, Clemson, SC 29634  
 Tel: 864 656 5737  
 Fax: 864 656 4293  
 Email: rdean@clemson.edu  
 Seq primer: TAATACGACTCTACTATAGGG

AQ361561 720 bp DNA linear GSS 03-FEB-1999  
 mgxb0004B19r CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0004B19r, genomic survey sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Magnaporthe grisea (anamorph: Pyricularia grisea)  
 Magnaporthe grisea  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.  
 1 (bases 1 to 720)  
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasnowski, M., Wing, R.A. and Dean, R.A.



**TITLE** A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome  
**JOURNAL** Unpublished (1998)  
**COMMENT** Contact: Dean RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson University, Clemson, SC 29634  
 Tel: 864 656 5737  
 Fax: 864 656 4293  
 Email: rdean@clemson.edu  
 Seq primer: GGAAACAGCTATGACCATG  
 Class: BAC ends  
 High quality sequence stop: 321.  
 Location/Qualifiers  
 1. .720

**FEATURES**  
 source  
 /organism="Magnaporthe grisea"  
 /mol\_type="genomic DNA"  
 /strain="70-15"  
 /db\_xref="taxon:148305"  
 /clone="mgxb0004B19r"  
 /tissue\_type="Protoplasts"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="CUGI Rice Blast BAC Library"  
 /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25x genome coverage. High density colony filters are available upon request."

**ORIGIN**

Query Match 12.4%; Score 82.2; DB 8; Length 720;  
 Best Local Similarity 54.1%; Pred. No. 2.5e-12;  
 Matches 236; Conservative 0; Mismatches 189; Indels 11; Gaps 3;  
 QY 148 AAGGTTAAAAATTCATGAAACACAAACACCAACCAAGTTGGTAACTATGTCCTATAAAC 207  
 DB 445 AAGGAAAAAAGAGAGACTAACCAACACCAACAAAAA-----CAGCCGCGTCATCAAC 391  
 QY 208 TAGGAGCCAACTTCCAAACCAATGGTAATGCGTATTTATGCGTCTATGTTGGACTGTT 267  
 DB 390 TACTCGGGCAGCTACAGCCNCGAGGCACTCATACCTGGCCGCTCTACGCGTGGACGCGC 331  
 QY 268 GACCTCTTTCGATATTTATTTGTCGACAGTTGGGCAACTGGCGTCCACCA---GGA 324  
 DB 330 AACCCGCTGATCGAGTACTACGTGGTGGAGAGCTTTGGCAGCTACAAACCGTCTGCGGC 271  
 QY 325 GCAACGCTTAAGGGGACCATCTGTTGATGGAGGAACATATGATATCTACGAGACTCTT 384  
 DB 270 GCAACCAACCGCGGCTCTTACCTCGACGGGAGCCTACGACATCTCTGTCAGCACC 211  
 QY 385 AGAGTCAATCAACCTCCATTAAGGGGATGCGCATTTAAACAATATGGAGTGTTCGA 444  
 DB 210 CGCTACAAACAGCCCTCCATCGACGGCACCAGAACCTTTACGAGTCTTGGTGGTGGC 151  
 QY 445 AGATCGAAACGACAGTGGCAGCATTTCTGTCAGCAACCACTTTAGAGCGTGGGAAC 504  
 DB 150 CGCAACAAGCGCGGCGGACCGCTCACCTTTGCCAACCAACGCTCAACGCTGGCGAAC 91  
 QY 505 TTAGGGATGAATATGGGGAATATGTAAGTTCGCG---CTTACTGTAGAAGGCTATCAA 561  
 DB 90 GCCGCGCTCACTCGGCAACAGTGGAACTACAGATCTTGGCGCTCGAGGGCTACCA 31  
 QY 562 AGTAGCGGAGTGCTA 577  
 DB 30 AGCAGCGGCTCGGCA 15

**RESULT 11**  
 AQL60254/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

**REFERENCE**  
 AUTHORS

**TITLE**  
**JOURNAL**  
**COMMENT**

**FEATURES**  
 source  
 /organism="Magnaporthe grisea"  
 /mol\_type="genomic DNA"  
 /strain="70-15"  
 /db\_xref="taxon:148305"  
 /clone="mgxb0003L19r"  
 /tissue\_type="Protoplasts"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="CUGI Rice Blast BAC Library"  
 /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25x genome coverage. High density colony filters are available upon request."

**ORIGIN**

Query Match 12.4%; Score 82.2; DB 8; Length 750;  
 Best Local Similarity 54.1%; Pred. No. 2.5e-12;  
 Matches 236; Conservative 0; Mismatches 189; Indels 11; Gaps 3;  
 QY 148 AAGGTTAAAAATTCATGAAACACAAACACCAACCAAGTTGGTAACTATGTCCTATAAAC 207  
 DB 445 AAGGAAAAAAGAGAGACTAACCAACACCAACAAAAA-----CAGCCGCGTCATCAAC 391  
 QY 208 TAGGAGCCAACTTCCAAACCAATGGTAATGCGTATTTATGCGTCTATGTTGGACTGTT 267  
 DB 390 TACTCGGGCAGCTACAGCCNCGAGGCACTCATACCTGGCCGCTCTACGCGTGGACGCGC 331  
 QY 268 GACCTCTTTCGATATTTATTTGTCGACAGTTGGGCAACTGGCGTCCACCA---GGA 324  
 DB 330 AACCCGCTGATCGAGTACTACGTGGTGGAGAGCTTTGGCAGCTACAAACCGTCTGCGGC 271  
 QY 325 GCAACGCTTAAGGGGACCATCTGTTGATGGAGGAACATATGATATCTACGAGACTCTT 384  
 DB 270 GCAACCAACCGCGGCTCTTACCTCGACGGGAGCCTACGACATCTCTGTCAGCACC 211  
 QY 385 AGAGTCAATCAACCTCCATTAAGGGGATGCGCATTTAAACAATATGGAGTGTTCGA 444  
 DB 210 CGCTACAAACAGCCCTCCATCGACGGCACCAGAACCTTTACGAGTCTTGGTGGTGGC 151  
 QY 445 AGATCGAAACGACAGTGGCAGCATTTCTGTCAGCAACCACTTTAGAGCGTGGGAAC 504  
 DB 150 CGCAACAAGCGCGGCGGACCGCTCACCTTTGCCAACCAACGCTCAACGCTGGCGAAC 91  
 QY 505 TTAGGGATGAATATGGGGAATATGTAAGTTCGCG---CTTACTGTAGAAGGCTATCAA 561  
 DB 90 GCCGCGCTCACTCGGCAACAGTGGAACTACAGATCTTGGCGCTCGAGGGCTACCA 31  
 QY 562 AGTAGCGGAGTGCTA 577  
 DB 30 AGCAGCGGCTCGGCA 15

Db 210 CGCTACACAGCCCTCCATCGACGGCACCAAGACCTTTACGAGTCTTGTCGGTGGCG 151  
 QY 445 AGATCGAAGCAGAGTGGCAGGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAAC 504  
 Db 150 CCACAACAGCGCCGAGCGGACCGCTCACTTTGCGCAACCAAGCTCAACGCTGGCGCAAC 91  
 QY 505 TTAGGATCAATATATGCGGAAATATGTAAGTTCGGG---CTTACTGTAGAGGCTATCAA 561  
 Db 90 GCGGCGCTCACTCGGCAACAGTGGAACTACGAGTCTTGCGCGTGCAGGGCTACCAAC 31  
 QY 562 AGTAGCGGAAGTGCTA 577  
 Db 30 AGCAGCGGCTCGGCCA 15

RESULT 12  
 AQ325248/c  
 LOCUS  
 DEFINITION  
 mgxb0021M08r CUGI Rice Blast BAC Library Magnaporthe grisea genomic  
 clone mgxb0021M08r, genomic survey sequence.  
 ACCESSION  
 AQ325248  
 VERSION  
 AQ325248.1 GI:4117100  
 KEYWORDS  
 GSS.  
 SOURCE  
 Magnaporthe grisea (anamorph: Pyricularia grisea)  
 ORGANISM  
 Magnaporthe grisea  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Sordariomycetes; incertae sedis; Magnaportheaceae; Magnaporthe.  
 1 (bases 1 to 786)  
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,  
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.  
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea  
 Genome

JOURNAL  
 COMMENT  
 Unpublished (1998)  
 Contact: Dean RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson University, Clemson, SC 29634  
 Tel: 864 656 5737  
 Fax: 864 656 4293  
 Email: rdean@clemson.edu  
 Seq primer: GGAAACAGCTATGACCATG  
 Class: BAC ends  
 High quality sequence start: 49  
 High quality sequence stop: 468.  
 Location/Qualifiers  
 1..786  
 /organism="Magnaporthe grisea"  
 /mol\_type="genomic DNA"  
 /strain="70-15"  
 /db\_xref="taxon:148305"  
 /clone="mgxb0021M08r"  
 /tissue\_type="Protoplasts"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="CUGI Rice Blast BAC Library"  
 /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;  
 Rice blast is one of the most devastating fungal diseases  
 of rice world wide. It is a filamentous ascomycete with  
 a haploid genome (n=7) of approximately 40 Mbp. Rice  
 blast is an important model fungal pathogen for studying  
 numerous aspects of the fungal-host interaction. In  
 order to facilitate genome wide analysis, a BAC library  
 containing 9216 clones with an average insert size of 130  
 kbp was constructed. This library represents greater  
 than 25X genome coverage. High density colony filters  
 are available upon request."

FEATURES  
 source

ORIGIN

Query Match 12.3%; Score 81.6; DB 8; Length 786;  
 Best Local Similarity 54.1%; Pred. No. 3 8e-12;  
 Matches 236; Conservative 0; Mismatches 189; Indels 11; Gaps 3;  
 QY 148 AAGGTGTAATAATTCATGAACACAAACACCAACAGTGGTAAACATGTCCATAAC 207  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 500 AAGGAAAAAAGAGACTAAACAAACAAACAAAAA-----CAGCCGGCTCATCAAC 446  
 QY 208 TACGAGCAACTTCCACCAAAATGTAATCGGTATTTATGCTCTATGTTGGACTGTT 267  
 Db 445 TACTCGGCGACTACAGCCCCAGGGCAACTCATACCTTGGCCGCTACGGGTGGACGGCG 386  
 QY 268 GACCTCTTGTGCAATATTATATTGTGACAGTTGGGGCAACTGGCGTCCACCA---GGA 324  
 Db 385 AACCGCTGTAGTACTAGTGTGTGAGAGCTTTTGGCAGCTACAACCCGCTGTCGGGCG 326  
 QY 325 GCAAAGCCTAAGGGACCAATCACTGTTGATGGAGAAACATATGATATCTACGAGACTTT 384  
 Db 325 GCCACCAACCGCGGTCCTTCACTCGGAGCGGACACTACGACATCCTGTCAGCAC 266  
 QY 385 AGAGTCAATCAACCTCCATTAAAGGGATTGCCACATTTAAACAATATTGAGTGTTCGA 444  
 Db 265 CGTACACACCAAGCCCTCCATCGACGGCACCAAGACCTTTACGAGTCTTGTCGGTGGCG 206  
 QY 445 AGATCGAAGCAGCAGGTGGCAGCATTTCTGTGACCAACCACTTTAGAGCGTGGGAAAC 504  
 Db 205 CGCAACAAGCGCGGCGGCGGACCGTCACCTTTGCCAACCAAGCTCAACGCTGGCGCAC 146  
 QY 505 TTAGGATGAATATGGGAAAAATGTATGAAGTCCG---CTTACTGTAGAAGGCTATCAA 561  
 Db 145 GCCGCGCTCAACCTCGGCAACCAAGTGAACCTACAGATCCTTGGCCGCTCGAGGGCTACCAC 86  
 QY 562 AGTAGCGGAAGTGCTA 577  
 Db 85 AGCAGCGGCTCGGCCA 70

RESULT 13

AQ399120/c

LOCUS

DEFINITION

mgxb0001B24f CUGI Rice Blast BAC Library Magnaporthe grisea genomic

clone mgxb0001B24f, genomic survey sequence.

ACCESSION

AQ399120

VERSION

AQ399120.1 GI:4370147

KEYWORDS

GSS.

SOURCE

Magnaporthe grisea (anamorph: Pyricularia grisea)

ORGANISM

Magnaporthe grisea

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

Sordariomycetes; incertae sedis; Magnaportheaceae; Magnaporthe.

1 (bases 1 to 583)

Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,

Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.

A BAC End Sequencing Framework to Sequence the Magnaporthe grisea

Genome

Unpublished (1998)

Contact: Dean RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson University, Clemson, SC 29634

Tel: 864 656 5737

Fax: 864 656 4293

Email: rdean@clemson.edu

Seq primer: TAATACGACTCACTATAGG

Class: BAC ends

High quality sequence stop: 411.

Location/Qualifiers

1..583

/organism="Magnaporthe grisea"

/mol\_type="genomic DNA"

/strain="70-15"

/db\_xref="taxon:148305"

/clone="mgxb0001B24f"

/tissue\_type="Protoplasts"

/lab\_host="E. coli DH10B"

/clone\_lib="CUGI Rice Blast BAC Library"

/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;  
 Rice blast is one of the most devastating fungal diseases  
 of rice world wide. It is a filamentous ascomycete with  
 a haploid genome (n=7) of approximately 40 Mbp. Rice

blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

## ORIGIN

Query Match	12.3%	Score 81.4;	DB 8;	Length 583;
Best Local Similarity	52.8%;	Pred. No. 4.1e-12;		
Matches 200;	Conservative 0;	Mismatches 176;	Indels 3;	Gaps 1;
Qy	154	AAAAATTCAATGAAACACAAACACACCAACAAGTTGGTAACATGTCCATAAACTACGGA	213	
Db	445	AAAAAAAAAGAGAGACTACATCNACAAACAACAAAACAGCCGCGTCATCAACTACTCG	386	
Qy	214	GCACACTTCCAAACAAATGGTAATGCGTATTTATGCGTCTATGTTGGAGCTGTTGACCC	273	
Db	385	GGCAGCTTACAGCCACACAGGGCAACTACACTCTGCGCGTCTACGCGTGGACGCGCAACCG	326	
Qy	274	CTTGTGCAATATTATTGTTCGACAGTTGGGGCAACTGGCGTCCACCA---GGAGCAACG	330	
Db	325	CTGATCAGTACTACGTTGGTGGAGAGCTTTGGCAGCTACACCCGTCGTGGGGCGCCACC	266	
Qy	331	CCTAAGGGGACCATCACTGTTGTATGGAGGAACATATGATATCTACGAGACTCTTAGAGTC	390	
Db	265	AACCGCGGTCCTTCACTCGGACGGCAGCACCTTACGACATCTCTGGTCAGCACCCGCTAC	206	
Qy	391	AATCAACCCCTCCATTAAAGGGATTGCCACATTTTAAACAATATTGGAGTGTTCGAGATCG	450	
Db	205	AACGAGCCCTCCATTCGACGGCAACAAACCCCTTTCAGCAGTTCTGGTCGGTGC	146	
Qy	451	AAAGCGACGAGTGGCAGATTCTTGTTCAGCAACCACTTTTAGCGCGTGGGAAAACTTAGGG	510	
Db	145	AAGCGGCCACGGCACCGTCACTTTTGGCCACACCGTCAACGCTGGCGCAACGCCGCG	86	
Qy	511	ATGAATATGGGAAAAATGT	529	
Db	85	CTCAACCTCGGCACCACT	67	

RESULT 14	561 bp	DNA	linear	GSS 06-MAR-1999
AQ396475/c	AQ396475	CUGI Rice Blast	EAC Library	Magnaporthe grisea genomic
LOCUS	mgxb0010M14f	clone maxb0010M14f	genomic survey	sequence.
DEFINITION				

AQ396475  
AO396475.1 GI:4367502

GSS.  
Magnaporthe grisea (anamorph: Pyricularia grisea)  
Magnaporthe grisea

magnaporthe grisea  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthae.

1 (bases 1 to 561)  
Yu, Y.; Zhu, H.; Boyd, C.A.; Gaudette, B.; Gayle, A.; Kingsbury, R.,

Phillips, K., Sasinowski, M, Wing, R.A. and Dean, R.A.  
A BAC End Sequencing Framework to Sequence the Magnaporthe grisea

Genome  
A PAC film sequencer

Unpublished (1998)

Contact: Dean RA

**Clemson University Genomics Institute**

Clemson University

100 Jordan Hall, Clemson University, Clemson, SC 29634

**Tel: 864 656 5737**

**Fax: 864 656 4293**

Email: [rdean@clemonson.edu](mailto:rdean@clemonson.edu)

Seq primer: TAATAGACTCACTATAGGG

**Class: BAC ends**

High quality sequence start: 50

High quality sequence stop: 443.

Location/Qualifiers

1. 561

```

/organism="Magnaporthe grisea"
/mol_type="genomic DNA"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0010M14f"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice Blast BAC Library"
/notes="vector: pBACWICH; Site_1: HindIII; Site_2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

```

## ORIGIN

Query Match	11.6%; Score 76.6; DB 8; Length 561;
Best Local Similarity	53.7%; Pred. No. 9.6e-11;
Matches	204; Conservative 0; Mismatches 170; Indels 6; Gaps 2;
QY	139 TTATTCGGTAAAGGTAAAAAAATTCATATGAACACAAAACACACCAACCAAGTTGGTAAACATG 198
Db	
QY	495 TTTTCCCAAAAAGGAAAAAAGAGAGACTANCAACACCAACAAAACAGCCGGCT- 437
Db	
QY	199 TCCATAAACTACGGAGCCAACTTCCAAACCAAAATGGTAAATGGTATTTATGGCTCTATGGT 258
Db	
QY	436 --CATCAACTACTCGGGCAGCTACAGCCCCCAGGGCAACTCATACCTGGCGCTCTACGGC 379
Db	
QY	259 TGGACTGTTGACCCCTCTGTTCGAATATATATGTCGCACAGTTGGGGCAACTGGCGTCCA 318
Db	
QY	378 TGGACGGCGCAACCCCGCTGATCGAGTACTACGTGTGGAGAGCTTTGGCAGCTACAACCCG 319
Db	
QY	319 CCA---GGAGCAACGCCCTAAAGGGGACCATCACTGTTGATGGAGAAACATATGATATCTAC 375
Db	
QY	318 TCGTGGGGGCGCCACCAACCGGGGTCTCTTCACTCGGAGGCGCAGACCTACGACATCCTG 259
Db	
QY	376 GAGACTCTTAGAGTCAATCAACCTTCCATTAAAGGGGATGGCCACATTTAAACAAATATGG 435
Db	
QY	258 GTCAGCACCCGCTACAAACCAAGCCCTCCATCGACGGCACCAAGACCTTTCAGCAGTTCTGG 199
Db	
QY	436 AGTCTTCGAAGATCGAAACGCAAGATGGGACGATTTCTGTGCAGCAACCCTTTAGAGCG 495
Db	
QY	198 TCGGTGGCGGCAACAGAGCGGCGCAGGGCACCCTGACCTTTGCCAACCAAGCTCAACGCC 139
Db	
QY	496 TGGAAAAACTTAGGGATGAA 515
Db	
QY	138 TGGCGCAACGCGCGCCTCAA 119
Db	

51. T. H. J. S. B.

AF638869	AF638869	617 bp	linear	EST 05-MAY-2000
LOCUS				
DEFINITION	MgC Mycosphaerella graminicola cDNA clone mgc12d03f, mRNA sequence.			
ACCESSION	AF638869			
VERSION	AF638869.1			
KEYWORDS	GI:47031926			
SOURCE	EST.			
ORGANISM	Mycosphaerella graminicola			
	Mycosphaerella graminicola			
	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et Chaetothyriomycetes incertae sedis; Mycosphaerellaceae;			
REFERENCE	Mycosphaerella.			
AUTHORS	1 (bases 1 to 617)			
TITLE	Keon, J.P.R., Hargreaves, J.A., Antoniw, J.F. and Hammond-Kosack, K. Analysis of expressed sequence tags from the wheat fungal leaf blotch pathogen, Mycosphaerella graminicola (anamorph Septoria tritici)			
JOURNAL	Fungal Genet. Biol. (2004) In press			

COMMENT

Contact: Keon J  
Plant Pathogen Interactions Division,  
Rothamsted Research,  
Harpenden, Herts, UNITED KINGDOM  
Tel: +44(0)1582 763133  
Fax: +44(0)1582 760981  
Email: john.keon@bbsrc.ac.uk  
Insert Length: 800 Std Error: 100.00  
Seq primer: M13 reverse.

FEATURES

Location/Qualifiers  
1..617  
/organism="Mycosphaerella graminicola"  
/mol\_type="mRNA"  
/strain="Strit"  
/db\_xref="taxon:54734"  
/clone="mgcl2d03f"  
/clone\_lib="Mgc"  
/note="Vector: pSPORT1; Library constructed from senescent  
wheat leaves 21-25 days after infection with  
Mycosphaerella graminicola exhibiting abundant hyphal  
growth and asexual sporulation"

ORIGIN

Query Match	11.5%;	Score 76.4;	DB 1;	Length 617;
Best Local Similarity	55.6%;	Pred. No. 1.1e-10;		
Matches 168;	Conservative 0;	Mismatches 131;	Indels 3;	Gaps 1;

  

QY	200	CCATAAACTACGAGGCCAATCTCCAAACCAATGGTAATCGGTATTTATCGGTCTATGGTT	259
Db	305	CCATCAACTACTCCGGCTCTTCAACCCCTCCGGCAACGGCTACATCGCGCTACGGCT	364
QY	260	GGACTGTTGACCCCTTGTGCAATATATATGTCGACAGTTGGGGCAACTGGCGTCCAC	319
Db	365	GGACCCGCAACCCATGTATCGAATACTACGTGTCGAATCTCTACGGCACCTTACAACCCCA	424
QY	320	---CAGGAGCAACGGCTAAGGGGACCATCAGTGTGATGGAGAAACATATGATATCTACG	376
Db	425	GCTCTGGGCTCAGAAGAGGGAACCGTCACCCAGCGGGGCACCTACGATATCTCTCC	484
QY	377	AGACTCTTAGTCAATCAACCTCCATTAAAGGGGATTCGCACATTTAAACAAATATTGA	436
Db	485	AAACCAACCCGCTGTCACACCAACGACCTCCATTGACGCACGACAGACCTTTTCAGCACTTCTGGA	544
QY	437	GTGTTGAGATCGAAACGACAGTGGCAGCATTTCTGTCAGCAACCACTTTAGAGCGT	496
Db	545	GGTGGCCAGCAGAAGCCGCTCGCGGCACCGTGACCATGAGAACCACTTCGACGCTC	604
QY	497	GG 498	
Db	605	GG 606	

Search completed: November 11, 2004, 01:16:47  
Job time : 2880.04 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 15:57:32 ; Search time 3599.04 Seconds  
(without alignments)  
9775.823 Million cell updates/sec

Title: US-09-909-207-4  
Perfect score: 744  
Sequence: 1 ATGACACAAAAGAAATTCAC.....TAACTTGGATAAAACAAT 744

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	744	100.0	744	6 A48225	A48225 Sequence 4
2	744	100.0	744	6 A48226	A48226 Sequence 5
3	744	100.0	744	6 AR193051	AR193051 Sequence
4	744	100.0	744	6 AR193052	AR193052 Sequence 10
5	744	100.0	1513	6 A48231	A48231 Sequence 10
6	744	100.0	1513	6 A48232	A48232 Sequence 11
7	744	100.0	1513	6 AR193055	AR193055 Sequence
8	744	100.0	1513	6 AR193056	AR193056 Sequence 1
9	713.6	95.9	744	6 A68006	A68006 Sequence 1
10	713.6	95.9	744	6 AR163110	AR163110 Sequence 18
11	671.6	90.3	744	6 A45313	A45313 Sequence 18
12	671.6	90.3	744	6 AR117325	AR117325 Sequence
13	663	89.1	663	6 A48222	A48222 Sequence 1
14	663	89.1	663	6 A48223	A48223 Sequence 2
15	663	89.1	663	6 AR193049	AR193049 Sequence
16	663	89.1	663	6 AR193050	AR193050 Sequence
17	587.2	78.9	744	6 A68016	A68016 Sequence 11
18	587.2	78.9	744	6 AR163117	AR163117 Sequence
19	355.8	47.8	2246	1 AB029319	AB029319 Bacillus

20	355.2	47.7	1531	1 BSXYNY	XS9059 Bacillus sp
21	281.8	37.9	1454	1 CLOXYNB	M31726 Clostridium
22	279.2	37.5	1070	1 BPXYNA	X00660 Bacillus pu
23	278.2	37.4	4128	1 AF047761	AF047761 Clostridi
24	277.6	37.3	687	1 AY526092	AY526092 Bacillus
25	276.6	37.2	5774	1 AB010958	AB010958 Clostridi
26	275	37.0	3493	1 CST508403	AJ508403 Clostridi
27	274.4	36.9	684	1 AF220528	AF220528 Bacillus
28	274.4	36.9	1011	1 AF490981	AF490981 Bacillus
29	273	36.7	1022	6 A42251	A42251 Sequence 1
30	273	36.7	1022	6 A42285	A42285 Sequence 35
31	273	36.7	1022	6 AR127019	AR127019 Sequence
32	273	36.7	1022	6 AR127049	AR127049 Sequence
33	273	36.7	1022	6 AR220023	AR220023 Sequence
34	273	36.7	1022	6 AR220053	AR220053 Sequence
35	273	36.7	1022	6 AR221306	AR221306 Sequence
36	273	36.7	1022	6 AR221336	AR221336 Sequence
37	272.8	36.7	681	6 A42280	A42280 Sequence 30
38	272.8	36.7	681	6 A42282	A42282 Sequence 32
39	272.8	36.7	681	6 AR127045	AR127045 Sequence
40	272.8	36.7	681	6 AR127046	AR127046 Sequence
41	272.8	36.7	681	6 AR220049	AR220049 Sequence
42	272.8	36.7	681	6 AR220050	AR220050 Sequence
43	272.8	36.7	681	6 AR221332	AR221332 Sequence
44	272.8	36.7	681	6 AR221333	AR221333 Sequence
45	269.6	36.2	1789	1 AF326785	AF326785 Bacillus

ALIGNMENTS

RESULT 1	A48225	Sequence 4 from Patent EP0698667.	744 bp	DNA	linear	PAT 07-MAR-1997
LOCUS	A48225	Sequence 4 from Patent EP0698667.				
DEFINITION	A48225	Sequence 4 from Patent EP0698667.				
ACCESSION	A48225	Sequence 4 from Patent EP0698667.				
VERSION	A48225.1	GI:2302072				
KEYWORDS	unidentified					
SOURCE	unidentified					
ORGANISM	unclassified.					
REFERENCE	1 (bases 1 to 744)					
AUTHORS	De,B.E., Lahaye,A., Ledoux,P. and Detroz,R.					
TITLE	Xylanase, microorganisms for its production, DNA molecules, process of preparation and use thereof					
JOURNAL	Patent: EP 0698667-A 4 28-FEB-1996;					
COMMENT	SOLVAY (BE)					
	Other publication BE 1008751 960702					
	Other publication BE 1008570 960604					
	Other publication BR 9503454 960305					
	Other publication JP 8092284 960409					
	Other publication FI 953578 960127					
	Other publication CA 2154628 960127					
	Other publication AU 2508695 960208.					

FEATURES	Location/Qualifiers
Source	1..744
	/organism="unidentified"
	/mol_type="unassigned DNA"
	/db_xref="taxon:32644"

ORIGIN

Query Match	100.0%;	Score 744;	DB 6;	Length 744;
Best Local Similarity	100.0%;	Pred. No. 1.9e-187;		
Matches 744;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGAGACAAAGAAATGACGTTAGCTTTTACCTTTTGTGTTGCTAACCTTA	60	
Db	1	ATGAGACAAAGAAATGACGTTAGCTTTTACCTTTTGTGTTGCTAACCTTA	60	
Qy	61	CCTGCGAATAATTCAGGCACAAATCGTCACCAATTCATTCGCAACACGATGGC	120	
Db	61	CCTGCGAATAATTCAGGCACAAATCGTCACCAATTCATTCGCAACACGATGGC	120	

```

QY 121 TATGATATGATTTTGGAAAGATAGCGGTGGCTCTGGGCAATGATTCTCAATCATGGC 180
Db 121 TATGATATGATTTTGGAAAGATAGCGGTGGCTCTGGGCAATGATTCTCAATCATGGC 180
QY 181 GGTACGTTCACTGCCCAATGAACAATGTTAAACAATATTTATTCGGTAAAGGTAAAAA 240
Db 181 GGTACGTTCACTGCCCAATGAACAATGTTAAACAATATTTATTCGGTAAAGGTAAAAA 240
QY 241 TTCAATGAAACACAAACACACAAAGTTGGTAAACATGTCATTAATCTAGGAGCAAC 300
Db 241 TTCAATGAAACACAAACACACAAAGTTGGTAAACATGTCATTAATCTAGGAGCAAC 300
QY 301 TTCCAAACCAATGTAATGCGTATTTATGCGCTCTATGTTGGATGTTGACCTCTTGTC 360
Db 301 TTCCAAACCAATGTAATGCGTATTTATGCGCTCTATGTTGGATGTTGACCTCTTGTC 360
QY 361 GAATATATATTTGTCAGAGTTGGGCAACTGGGCTCCACGAGGCAACGCCCTAAGGG 420
Db 361 GAATATATATTTGTCAGAGTTGGGCAACTGGGCTCCACGAGGCAACGCCCTAAGGG 420
QY 421 ACCATCACTGTTGATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACC 480
Db 421 ACCATCACTGTTGATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACC 480
QY 481 TCCATTAGGGGATGCGCACATTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACG 540
Db 481 TCCATTAGGGGATGCGCACATTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACG 540
QY 541 AGTGCACGATTTCTGTGACGACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATG 600
Db 541 AGTGCACGATTTCTGTGACGACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATG 600
QY 601 GGGAAATATGTAAGTGGCTTACTGTAGAGGCTATCAAGTAGCGGAAGTCTTAAT 660
Db 601 GGGAAATATGTAAGTGGCTTACTGTAGAGGCTATCAAGTAGCGGAAGTCTTAAT 660
QY 661 GTATATAGCAATPACACTAAGATTAACGGTAAACCGTAAACCTCTCTCACTTAGTAGCAG 720
Db 661 GTATATAGCAATPACACTAAGATTAACGGTAAACCGTAAACCTCTCTCACTTAGTAGCAG 720
QY 721 AGCATAACTTTGGATAAAAAAAT 744
Db 721 AGCATAACTTTGGATAAAAAAAT 744

```

```

RESULT 2
A48226 744 bp DNA linear PAT 07-MAR-1997
LOCUS Sequence 5 from Patent EP0698667.
DEFINITION A48226
ACCESSION A48226
VERSION A48226.1 GI:2302073
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 744)
AUTHORS De,B.E., Lahaye,A., Ledoux,P. and Detroz,R.
TITLE Xylanase, microorganisms for its production, DNA molecules, process
of preparation and use thereof
JOURNAL Patent: EP 0698667-A 5 28-FEB-1996;
SOLVAY (BE)
COMMENT Other publication BE 1008751 960702
Other publication BE 1008570 960604
Other publication BR 9503454 960305
Other publication JP 8052284 960409
Other publication FI 953578 960127
Other publication CA 2154628 960127
Other publication AU 2508695 960208.
FEATURES
Location/Qualifiers
1..744
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

```

```

CDS
1..>744
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA03092.1"
/db_xref="GI:2302074"
/translation="MRQKKLTLLFLVCFALTLPALIIQAIQIVTDSNGHGDYDYE
FKWDSGSGTMIINHGGTFSQOWNVNNILFRKKFNETQTHQQVGNMSINYGANFQ
PNGNAVLCVYGVMTVDPLVEYIVDSWGNRPPGATPKGTITVDGGYVDIYETURVNP
SINGIATFQYMSVRSRSTSGTISVSNHRAWENLGMNMGKYEVALIVEGYQSSGS
ANYSNLTRINGNPLSTIISNDESITLIDKNN"
sig_peptide 1..81
mat_peptide 82..744
/product="unnamed"
ORIGIN
Query Match 100.0%; Score 744; DB 6; Length 744;
Best Local Similarity 100.0%; Pred. No. 1.9e-187;
Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGACAAAGAAATTTGACGTTGATTTTAGCTTTTAGCTTTTGTGTTGCACTAACCTTA 60
Db 1 ATGAGACAAAGAAATTTGACGTTGATTTTAGCTTTTGTGTTGCACTAACCTTA 60
QY 61 CCTGCAGAAATAATTCAGGCACAAATCGTCCAGCAAAATCCATTTGGCAACCCAGATGGC 120
Db 61 CCTGCAGAAATAATTCAGGCACAAATCGTCCAGCAAAATCCATTTGGCAACCCAGATGGC 120
QY 121 TATGATATCAATTTTGGAAAGATAGCGGTGGCTCTGGGCAATGATTCTCAATCATGGC 180
Db 121 TATGATATCAATTTTGGAAAGATAGCGGTGGCTCTGGGCAATGATTCTCAATCATGGC 180
QY 181 GGTACGTTCACTGCCCAATGAACAATGTTAAACAATATTTATTCGTAAGGTAAAAA 240
Db 181 GGTACGTTCACTGCCCAATGAACAATGTTAAACAATATTTATTCGTAAGGTAAAAA 240
QY 241 TTCAATGAAACACAAACACACCAAGTTGGTAAACATGTCATTAATCTAGGAGCAAC 300
Db 241 TTCAATGAAACACAAACACACCAAGTTGGTAAACATGTCATTAATCTAGGAGCAAC 300
QY 301 TTCCAAACCAATGTAATGCGTATTTATGCGCTCTATGTTGGATGTTGACCTCTTGTC 360
Db 301 TTCCAAACCAATGTAATGCGTATTTATGCGCTCTATGTTGGATGTTGACCTCTTGTC 360
QY 361 GAATATATATTTGTCAGAGTTGGGCAACTGGGCTCCACGAGGCAACGCCCTAAGGG 420
Db 361 GAATATATATTTGTCAGAGTTGGGCAACTGGGCTCCACGAGGCAACGCCCTAAGGG 420
QY 421 ACCATCACTGTTGATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACC 480
Db 421 ACCATCACTGTTGATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACC 480
QY 481 TCCATTAGGGGATGCGCACATTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACG 540
Db 481 TCCATTAGGGGATGCGCACATTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACG 540
QY 541 AGTGCACGATTTCTGTGACGACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATG 600
Db 541 AGTGCACGATTTCTGTGACGACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATG 600
QY 601 GGGAAATATGTAAGTGGCTTACTGTAGAGGCTATCAAGTAGCGGAAGTCTTAAT 660
Db 601 GGGAAATATGTAAGTGGCTTACTGTAGAGGCTATCAAGTAGCGGAAGTCTTAAT 660
QY 661 GTATATAGCAATACACTAAGATTAACGGTAAACCGTAAACCTCTCTCACTTAGTAGCAG 720
Db 661 GTATATAGCAATACACTAAGATTAACGGTAAACCGTAAACCTCTCTCACTTAGTAGCAG 720
QY 721 AGCATAACTTTGGATAAAAAAAT 744
Db 721 AGCATAACTTTGGATAAAAAAAT 744

```

```
AR193051
LOCUS AR193051 744 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 4 from patent US 6346407.
ACCESSION AR193051
VERSION AR193051.1 GI:20239016
KEYWORDS
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 744)
  De Buyl, E., Lahaye, A., Ledoux, P. and Detroz, R.
  Xylanase, microorganisms producing it, DNA molecules, methods for
  preparing this xylanase and uses of the latter
  Patent: US 6346407-A 4 12-FEB-2002;
  Location/Qualifiers
  1..744
  /organism="unknown"
  /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 744; DB 6; Length 744;
Best Local Similarity 100.0%; Pred. No. 1.9e-187;
Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGCAAAAGAAATTGACGTTGATTTAGCCTTTTGTAGTTGTTTGCACCTTA 60
DB 1 ATGAGCAAAAGAAATTGACGTTGATTTAGCCTTTTGTAGTTGTTTGCACCTTA 60
QY 61 CTGCGAGAAATTAATTCAGGCACAAATCGTCACGCAAAATTCATTTGGCAACACGATGC 120
DB 61 CTGCGAGAAATTAATTCAGGCACAAATCGTCACGCAAAATTCATTTGGCAACACGATGC 120
QY 121 TATGATTATGAATTTGGAAAGATAGCGGTGGCTCTGGGCAATGATTTCTCAATCATGCG 180
DB 121 TATGATTATGAATTTGGAAAGATAGCGGTGGCTCTGGGCAATGATTTCTCAATCATGCG 180
QY 181 GGTACGTTTCAGTCCCAATGGAACAATGTTAAACAATATTTCCGTTAAAGTAAATAA 240
DB 181 GGTACGTTTCAGTCCCAATGGAACAATGTTAAACAATATTTCCGTTAAAGTAAATAA 240
QY 241 TTCAATGAAACACAAACACACCAACCAAGTTGGTAAACATGTTCCATAACTACGAGCCAAC 300
DB 241 TTCAATGAAACACAAACACACCAACCAAGTTGGTAAACATGTTCCATAACTACGAGCCAAC 300
QY 301 TTCCAAACCAATGTTGATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 360
DB 301 TTCCAAACCAATGTTGATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 360
QY 361 GAATATTTATTTGTCACAGTTGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGGG 420
DB 361 GAATATTTATTTGTCACAGTTGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGGG 420
QY 421 ACCATCACTGTTGATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480
DB 421 ACCATCACTGTTGATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480
QY 481 TCCATTAAAGGGGATTCGCCAATTTAAACAATATTTGGAGTGTTCGAAGATCGAAGCGCAG 540
DB 481 TCCATTAAAGGGGATTCGCCAATTTAAACAATATTTGGAGTGTTCGAAGATCGAAGCGCAG 540
QY 541 AGTGGCACGATTTCTGTGACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 600
DB 541 AGTGGCACGATTTCTGTGACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 600
QY 601 GGGAAATATGTAAGTCGCGCTTACTGTAGAGGGCTATCAAAAGTAGCGGAAGTCTTAAT 660
DB 601 GGGAAATATGTAAGTCGCGCTTACTGTAGAGGGCTATCAAAAGTAGCGGAAGTCTTAAT 660
QY 661 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTAATGACGAG 720
DB 661 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTAATGACGAG 720
QY 721 AGCATAACTTTGGTAAACAAAT 744
```

```
DB 721 AGCATAACTTTGGTAAACAAAT 744
RESULT 4
LOCUS AR193052
DEFINITION Sequence 5 from patent US 6346407.
ACCESSION AR193052
VERSION AR193052.1 GI:20239017
KEYWORDS
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 744)
  De Buyl, E., Lahaye, A., Ledoux, P. and Detroz, R.
  Xylanase, microorganisms producing it, DNA molecules, methods for
  preparing this xylanase and uses of the latter
  Patent: US 6346407-A 5 12-FEB-2002;
  Location/Qualifiers
  1..744
  /organism="unknown"
  /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 744; DB 6; Length 744;
Best Local Similarity 100.0%; Pred. No. 1.9e-187;
Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGCAAAAGAAATTGACGTTGATTTAGCCTTTTGTAGTTGTTTGCACCTTA 60
DB 1 ATGAGCAAAAGAAATTGACGTTGATTTAGCCTTTTGTAGTTGTTTGCACCTTA 60
QY 61 CTGCGAGAAATTAATTCAGGCACAAATCGTCACGCAAAATTCATTTGGCAACACGATGC 120
DB 61 CTGCGAGAAATTAATTCAGGCACAAATCGTCACGCAAAATTCATTTGGCAACACGATGC 120
QY 121 TATGATTATGAATTTGGAAAGATAGCGGTGGCTCTGGGCAATGATTTCTCAATCATGCG 180
DB 121 TATGATTATGAATTTGGAAAGATAGCGGTGGCTCTGGGCAATGATTTCTCAATCATGCG 180
QY 181 GGTACGTTTCAGTCCCAATGGAACAATGTTAAACAATATTTCCGTTAAAGTAAATAA 240
DB 181 GGTACGTTTCAGTCCCAATGGAACAATGTTAAACAATATTTCCGTTAAAGTAAATAA 240
QY 241 TTCAATGAAACACAAACACACCAACCAAGTTGGTAAACATGTTCCATAACTACGAGCCAAC 300
DB 241 TTCAATGAAACACAAACACACCAACCAAGTTGGTAAACATGTTCCATAACTACGAGCCAAC 300
QY 301 TTCCAAACCAATGTTGATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 360
DB 301 TTCCAAACCAATGTTGATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 360
QY 361 GAATATTTATTTGTCACAGTTGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGGG 420
DB 361 GAATATTTATTTGTCACAGTTGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGGG 420
QY 421 ACCATCACTGTTGATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480
DB 421 ACCATCACTGTTGATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480
QY 481 TCCATTAAAGGGGATTCGCCAATTTAAACAATATTTGGAGTGTTCGAAGATCGAAGCGCAG 540
DB 481 TCCATTAAAGGGGATTCGCCAATTTAAACAATATTTGGAGTGTTCGAAGATCGAAGCGCAG 540
QY 541 AGTGGCACGATTTCTGTGACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 600
DB 541 AGTGGCACGATTTCTGTGACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 600
QY 601 GGGAAATATGTAAGTCGCGCTTACTGTAGAGGGCTATCAAAAGTAGCGGAAGTCTTAAT 660
DB 601 GGGAAATATGTAAGTCGCGCTTACTGTAGAGGGCTATCAAAAGTAGCGGAAGTCTTAAT 660
QY 661 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTAATGACGAG 720
DB 661 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTAATGACGAG 720
QY 721 AGCATAACTTTGGTAAACAAAT 744
```





QY 61 CCTGCAGAAATTAATTCAGGCAAAATTCGTCAGGCAAAATTCATTTGGCAACCAACGATGCG 120  
DB 680 CCTGCAGAAATTAATTCAGGCAAAATTCGTCAGGCAAAATTCATTTGGCAACCAACGATGCG 739  
QY 121 TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTGGGACAAATGATTTCTCAATCATGCG 180  
DB 740 TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTGGGACAAATGATTTCTCAATCATGCG 799  
QY 181 GGTACGTTTCAGTCCCAATGGAACAAATGTTTAAACAATATTTATTCGTTAAAGGTAATAAA 240  
DB 800 GGTACGTTTCAGTCCCAATGGAACAAATGTTTAAACAATATTTATTCGTTAAAGGTAATAAA 859  
QY 241 TTCAATGAAACACAAACACCAACCAAGTTGGTAAACATGTTTAAACAATATTTATTCGTTAAAGGTAATAAA 300  
DB 860 TTCAATGAAACACAAACACCAACCAAGTTGGTAAACATGTTTAAACAATATTTATTCGTTAAAGGTAATAAA 919  
QY 301 TTCCAAACCAATGTTTAAAGGTAATAAGTGGTAAACATGTTTAAACAATATTTATTCGTTAAAGGTAATAAA 360  
DB 920 TTCCAAACCAATGTTTAAAGGTAATAAGTGGTAAACATGTTTAAACAATATTTATTCGTTAAAGGTAATAAA 979  
QY 361 GAATATTATTTATGTCAGCAAGTTGGGCAACCTGGCGTCCACAGGAGCAACGGCTTAAGGGG 420  
DB 980 GAATATTATTTATGTCAGCAAGTTGGGCAACCTGGCGTCCACAGGAGCAACGGCTTAAGGGG 1039  
QY 421 ACCATCACTGTTGATGGAGCAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480  
DB 1040 ACCATCACTGTTGATGGAGCAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 1099  
QY 481 TCCATTAAAGGGGATTCGCAATTTAAACAATATTTAGAGTGTTCGAAAGTCAATCAACCC 540  
DB 1100 TCCATTAAAGGGGATTCGCAATTTAAACAATATTTAGAGTGTTCGAAAGTCAATCAACCC 1159  
QY 541 AGTGGCAAGGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 600  
DB 1160 AGTGGCAAGGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 1219  
QY 601 GGGAAATGATGAAGTTCGCGCTTACTGTAGAGGCTTACAAAGTACGCGAAGTGTCTAAT 660  
DB 1220 GGGAAATGATGAAGTTCGCGCTTACTGTAGAGGCTTACAAAGTACGCGAAGTGTCTAAT 1279  
QY 661 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTAATGACGAG 720  
DB 1280 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTAATGACGAG 1339  
QY 721 AGCATAACTTTGGATAAAACAAT 744  
DB 1340 AGCATAACTTTGGATAAAACAAT 1363

RESULT 7  
AR193055  
LOCUS AR193055 1513 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 10 from patent US 6346407.  
ACCESSION AR193055  
VERSION AR193055.1 GI:20239020  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1513)  
AUTHORS De Buyl, E., Lahaye, A., Ledoux, P. and Detroz, R.  
TITLE Xylanase, microorganisms producing it, DNA molecules, methods for preparing this xylanase and uses of the latter  
JOURNAL Patent: US 6346407-A 10 12-FEB-2002;  
FEATURES Location/Qualifiers  
source 1. .1513  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 100.0%; Score 744; DB 6; Length 1513;  
Best Local Similarity 100.0%; Pred. No. 1.8e-187;

Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGACACAAAAGAAATTCAGCTGTTGATTTAGCTTTTGTAGTTTGTGCACTAACCTTTA 60  
DB 620 ATGACACAAAAGAAATTCAGCTGTTGATTTAGCTTTTGTAGTTTGTGCACTAACCTTTA 679  
QY 61 CCTGCAGAAATTAATTCAGGCAAAATTCGTCAGGCAAAATTCATTTGGCAACCAACGATGCG 120  
DB 680 CCTGCAGAAATTAATTCAGGCAAAATTCGTCAGGCAAAATTCATTTGGCAACCAACGATGCG 739  
QY 121 TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTGGGACAAATGATTTCTCAATCATGCG 180  
DB 740 TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTGGGACAAATGATTTCTCAATCATGCG 799  
QY 181 GGTACGTTTCAGTCCCAATGGAACAAATGTTTAAACAATATTTATTCGTTAAAGGTAATAAA 240  
DB 800 GGTACGTTTCAGTCCCAATGGAACAAATGTTTAAACAATATTTATTCGTTAAAGGTAATAAA 859  
QY 241 TTCAATGAAACACAAACACCAACCAAGTTGGTAAACATGTTTAAACAATATTTATTCGTTAAAGGTAATAAA 300  
DB 860 TTCAATGAAACACAAACACCAACCAAGTTGGTAAACATGTTTAAACAATATTTATTCGTTAAAGGTAATAAA 919  
QY 301 TTCCAAACCAATGTTTAAAGGTAATAAGTGGTAAACATGTTTAAACAATATTTATTCGTTAAAGGTAATAAA 360  
DB 920 TTCCAAACCAATGTTTAAAGGTAATAAGTGGTAAACATGTTTAAACAATATTTATTCGTTAAAGGTAATAAA 979  
QY 361 GAATATTATTTATGTCAGCAAGTTGGGCAACCTGGCGTCCACAGGAGCAACGGCTTAAGGGG 420  
DB 980 GAATATTATTTATGTCAGCAAGTTGGGCAACCTGGCGTCCACAGGAGCAACGGCTTAAGGGG 1039  
QY 421 ACCATCACTGTTGATGGAGCAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480  
DB 1040 ACCATCACTGTTGATGGAGCAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 1099  
QY 481 TCCATTAAAGGGGATTCGCAATTTAAACAATATTTAGAGTGTTCGAAAGTCAATCAACCC 540  
DB 1100 TCCATTAAAGGGGATTCGCAATTTAAACAATATTTAGAGTGTTCGAAAGTCAATCAACCC 1159  
QY 541 AGTGGCAAGGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 600  
DB 1160 AGTGGCAAGGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 1219  
QY 601 GGGAAATGATGAAGTTCGCGCTTACTGTAGAGGCTTACAAAGTACGCGAAGTGTCTAAT 660  
DB 1220 GGGAAATGATGAAGTTCGCGCTTACTGTAGAGGCTTACAAAGTACGCGAAGTGTCTAAT 1279  
QY 661 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTAATGACGAG 720  
DB 1280 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTAATGACGAG 1339  
QY 721 AGCATAACTTTGGATAAAACAAT 744  
DB 1340 AGCATAACTTTGGATAAAACAAT 1363

RESULT 8  
AR193056  
LOCUS AR193056 1513 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 11 from patent US 6346407.  
ACCESSION AR193056  
VERSION AR193056.1 GI:20239021  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1513)  
AUTHORS De Buyl, E., Lahaye, A., Ledoux, P. and Detroz, R.  
TITLE Xylanase, microorganisms producing it, DNA molecules, methods for preparing this xylanase and uses of the latter  
JOURNAL Patent: US 6346407-A 11 12-FEB-2002;  
FEATURES Location/Qualifiers  
source 1. .1513  
/organism="unknown"



[illegible]

```
Db 238 TTCAATGAAACACAAACACACCAACCAAGTTGGTAACTATGTCATTAACAACTATGGCGCAAC 297
Qy 301 TTCCAAACCAATGGTAATGCGTATTTATGCGTCTATGTTGGACTGTTGACCCCTCTTGTC 360
Db 298 TTCCAGCCAAACGGTAATGCGTATTTATGCGTCTATGTTGGACTGTTGACCCCTCTTGTT 357
Qy 361 GAAATATTATATTGTCAGAGTTGGGCAACTGGCGTCCACGAGGAGCAACGCCCTAAAGGG 420
Db 358 GAAATATTATATTGTCAGAGTTGGGCAACTGGCGTCCACGAGGAGCAACGCCCTAAAGGG 417
Qy 421 ACCATCACTGTTGATGGAGAACATATGATATCTACGAGCTCTTAGAGTCAATCAACCC 480
Db 418 ACCATCACTGTTGATGGAGAACATATGATATCTACGAGCTCTTAGAGTCAATCAACCC 477
Qy 481 TCCATTAAGGGATGCGCAATTTAAACAATATTGGAGTGTTCGAAGATCGAAACGCAAC 540
Db 478 TCCATTAAGGGATGCGCAATTTAAACAATATTGGAGTGTTCGAAGATCGAAACGCAAC 537
Qy 541 AGTGCACGATTTCTGTGACGACCTTTTAGAGCGTGGGAAACTTAGGGATGAATATG 600
Db 538 AGTGCCACAATTTCTGTGACGACCTTTTAGAGCGTGGGAAACTTAGGGATGAATATG 597
Qy 601 GGGAAATGTATGAAGTCGCGCTTACTGTAGAGGCTATCAAGCTAGCGGAAGTCTAAT 660
Db 598 GGGAAATGTATGAAGTCGCGCTTACTGTAGAGGCTATCAAGCTAGCGGAAGTCTAAT 657
Qy 661 GTATATAGCAATACACTAAGAAATTAACGGAACCCCTCTCAACTATTAGTAATACGAG 720
Db 658 GTATATAGCAATACACTAAGAAATTAACGGAACCCCTCTCAACTATTAGTAATACGAG 717
Qy 721 AGCATAACTTTGGATAAAAAACAAT 744
Db 718 AGCATAACTTAGATAAAAAACAAT 741

RESULT 12
AR117325 AR117325 744 bp DNA linear PAT 16-MAY-2001
LOCUS Sequence 18 from patent US 6140095.
ACCESSION AR117325
VERSION AR117325.1 GI:14098231
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 744)
AUTHORS Williams, D.P., Iverson, S., Farrell, R. Lee., Van Solingen, P.,
Herbes, W. Theresia., Van Der Kleij, W. Antonius H., Van
Beckhoven, R. Franciscus. C., Quax, W. Johannes., Herwijer, M. Adriana.,
Goedegebuur, F. and Jones, B. Edward.
TITLE Alkalitolerant xylanases
JOURNAL Patent: US 6140095-A 18 31-OCT-2000;
FEATURES Location/Qualifiers
source 1..744
/mol_type="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 90.3%; Score 671.6; DB 6; Length 744;
Best Local Similarity 95.7%; Pred. No. 3.9e-168;
Matches 712; Conservative 0; Mismatches 29; Indels 3; Gaps 2;

Qy 1 ATGACAAAAGAAATGACGTTGATTTAGCCTTTTGTAGTTTGTGTTGCACTAACCTTA 60
Db 1 ATGACAAAAGAAATGACGTTGATTTAGCCTTTTGTAGTTTGTGTTGCACTAACCTTA 57
Qy 61 CTTGCAGAAATTAATTCAGGCAAAATCGTCACCGCAATTCATTTGGCAACCAACGATGCG 120
Db 58 CTTGCAGAAATTAATTCAGGCAAAATCGTCACCGCAATTCATTTGGCAACCAACGATGCG 117
Qy 121 TATGATTATGAAATTTTGGAAAGATAGCGGTGGCTCTGGACATGATTTCTCAATCATGCG 180
Db 118 TATGATTATGAAATTTTGGAAAGATAGCGGTGGCTCTGGACATGATTTCTCAATCATGCG 177
```

```
Qy 181 GGTCAGTTTCAGTCCCAATGGAACAATGTTTAAACAATATTTATTCGCTAAAGGTAAAAAA 240
Db 178 GGTCAGTTTCAGTCCCAATGGAACAATGTTTAAACAATATTTATTCGCTAAAGGTAAAAAA 237
Qy 241 TTCAATGAAACACAAACACACCAACCAAGTTGGTAAACATGTCATATAAATACTATGGCGCAAC 300
Db 238 TTCAATGAAACACAAACACACCAACCAAGTTGGTAAACATGTCATATAAATACTATGGCGCAAC 297
Qy 301 TTCCAAACCAATGGTAATGCGTATTTATGCGTCTATGTTGGACTGTTGACCCCTCTTGTC 360
Db 298 TTCCAGCCAAACGGTAATGCGTATTTATGCGTCTATGTTGGACTGTTGACCCCTCTTGTT 357
Qy 361 GAAATATTATATTGTCAGAGTTGGGCAACTGGCGTCCACGAGGAGCAACGCCCTAAAGGG 420
Db 358 GAAATATTATATTGTCAGAGTTGGGCAACTGGCGTCCACGAGGAGCAACGCCCTAAAGGG 417
Qy 421 ACCATCACTGTTGATGGAGAACATATGATATCTACGAGCTCTTAGAGTCAATCAACCC 480
Db 418 ACCATCACTGTTGATGGAGAACATATGATATCTACGAGCTCTTAGAGTCAATCAACCC 477
Qy 481 TCCATTAAGGGATGCGCAATTTAAACAATATTGGAGTGTTCGAAGATCGAAACGCAAC 540
Db 478 TCCATTAAGGGATGCGCAATTTAAACAATATTGGAGTGTTCGAAGATCGAAACGCAAC 537
Qy 541 AGTGCACGATTTCTGTGACGACCTTTTAGAGCGTGGGAAACTTAGGGATGAATATG 600
Db 538 AGTGCCACAATTTCTGTGACGACCTTTTAGAGCGTGGGAAACTTAGGGATGAATATG 597
Qy 601 GGGAAATGTATGAAGTCGCGCTTACTGTAGAGGCTATCAAGCTAGCGGAAGTCTAAT 660
Db 598 GGGAAATGTATGAAGTCGCGCTTACTGTAGAGGCTATCAAGCTAGCGGAAGTCTAAT 657
Qy 661 GTATATAGCAATACACTAAGAAATTAACGGAACCCCTCTCAACTATTAGTAATACGAG 720
Db 658 GTATATAGCAATACACTAAGAAATTAACGGAACCCCTCTCAACTATTAGTAATACGAG 717
Qy 721 AGCATAACTTTGGATAAAAAACAAT 744
Db 718 AGCATAACTTAGATAAAAAACAAT 741

RESULT 13
AR48222 AR48222 663 bp DNA linear PAT 07-MAR-1997
LOCUS Sequence 1 from Patent EP0698667.
ACCESSION AR48222
VERSION AR48222.1 GI:2302069
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 663)
AUTHORS De, B. E., Lahaye, A., Ledoux, P. and Detroz, R.
TITLE Xylanase, microorganisms for its production, DNA molecules, process
of preparation and use thereof
JOURNAL Patent: EP 0698667-A 1 28-FEB-1996;
COMMENT SOLVAY (BE)
Other publication BE 1008751 960702
Other publication BE 1008570 960604
Other publication BR 9503454 960305
Other publication JP 8092284 960409
Other publication FI 953578 960127
Other publication CA 2154628 960127
Other publication AU 2508595 960208.
FEATURES Location/Qualifiers
source 1..663
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 89.1%; Score 663; DB 6; Length 663;
```

[illegible]

RESULT	14
A48223	
LOCUS	
DEFINITION	663 bp DNA linear PAT 07-MAR-1997
ACCESSION	Sequence 2 from Patent EP0698667.
A48223	
ACCESSION	
A48223.1	GI:2302070
VERSION	
KEYWORDS	
SOURCE	unidentified
ORGANISM	unidentified unclassified.
REFERENCE	1 (bases 1 to 663)
AUTHORS	De,B.E., Lahaye,A., Ledoux,P. and Petroz,R. Xylanase, microorganisms for its production, DNA molecules, process of preparation and use thereof
TITLE	Patent: EP 0698667-A 2 28-FEB-1996;
JOURNAL	SOLVAY (BE) Other publication BE 1008751 960702 Other publication BE 1008570 960604 Other publication BR 9503454 960305 Other publication JP 8092284 960409 Other publication FI 953578 960127
COMMENT	

```

Other publication CA 2154628 960127
Other publication AU 2508695 960208.
Location/Qualifiers
    1..663
        /organism="unidentified"
        /mol_type="unassigned DNA"
        /db_xref="taxon:32644"
    1..663
        /note="unnamed protein product; Protein sequence is in
        conflict with the conceptual translation"
        /codon_start=1
        /protein_id="CAA03091.1"
        /db_xref="GI:2302071"
        /translation="QIVTDSNGHDGYDFEWKDGGSGTMLNHGGTFSACQNNVN
        NIDFGKGFENFTQHQVCNNSIANGFANGNAYLCVGTWDPVLEYIVDSWG
        NNRPGATPGKTIIVDGGIYETILRNQPSIKGIATPKQYWSRSTSGTISVS
        KNHFWENLGMNMGMKMEYVALIVEGQSSGANVYSNLTIRLNGPLSTISNDESIIL
        KNN"
mat_peptide
    1..663
        /product="unnamed"

```

mat peptide

## ORIGIN

Query Match	89.1%	Score 663	DB 6	Length 663	
Best Local Similarity	100.0%	Pred. No. 7.7e-166			
Matches	663	Conservative 0	Mismatches 0	Indels 0	
Gaps	0				
QY	82	CAAAATCGTCACCGACAAATTC	CAATTTGGCCAAACACGATGGCTATGATTAATGAATTTTGAAA	141	
DB	1	CAAAATCGTCACCGACAAATTC	CAATTTGGCCAAACACGATGGCTATGATTAATGAATTTTGAAA	60	
QY	142	GATAGCGGTGGCTCTCGGACAATG	TCTCAATCATGGCGGTACGTTCACTGTCGCCAATGG	201	
DB	61	GATAGCGGTGGCTCTCGGACAATG	TCTCAATCATGGCGGTACGTTCACTGTCGCCAATGG	120	
QY	202	AACAATGTTAAACAATATATTT	TCCTCGGTAAAGGTAAAAAATTC	CAATGAACACAAACACAC	261
DB	121	AACAATGTTAAACAATATATTT	TCCTCGGTAAAGGTAAAAAATTC	CAATGAACACAAACACAC	180
QY	262	CAACAAGTTCGTAAACATGTC	CCATAAATCTACGGAGCCAACTT	CCAAACAAATGGTAATGCG	321
DB	181	CAACAAGTTCGTAAACATGTC	CCATAAATCTACGGAGCCAACTT	CCAAACAAATGGTAATGCG	240
QY	322	TATTTATGCGTCTATGGTTGG	ACTGTTGACCCCTCTTCGAAATATATAT	TGTCGACAGT	381
DB	241	TATTTATGCGTCTATGGTTGG	ACTGTTGACCCCTCTTCGAAATATATAT	TGTCGACAGT	300
QY	382	TGGGGCAACTGGCGTCCAC	GAGNACACCGCTTAAGGGGACCATCAT	CTGTTGATGGAGGA	441
DB	301	TGGGGCAACTGGCGTCCAC	GAGNACACCGCTTAAGGGGACCATCAT	CTGTTGATGGAGGA	360
QY	442	ACATATGATATCTACGAGACT	CTTTAGAGTCAATCAACCCCTCCAT	TAAAGGGGATTGCCACA	501
DB	361	ACATATGATATCTACGAGACT	CTTTAGAGTCAATCAACCCCTCCAT	TAAAGGGGATTGCCACA	420
QY	502	TTTTAAACAATATGGAGTGT	TCGAAGATCGAAACGCGACGAGTGGCACGAT	TTCTGTCTACG	561
DB	421	TTTTAAACAATATGGAGTGT	TCGAAGATCGAAACGCGACGAGTGGCACGAT	TTCTGTCTACG	480
QY	562	AACACATTTAGAGCGTGG	AAAACTTAGGAGTGAATATGGGGAAAAAT	TGTGAAGTCCCG	621
DB	481	AACACATTTAGAGCGTGG	AAAACTTAGGAGTGAATATGGGGAAAAAT	TGTGAAGTCCCG	540
QY	622	CTTACTGTAGAAGCTATCA	AGGTAGCGGAAGTCTAATGTATATAGCAATACACTAAGA		681
DB	541	CTTACTGTAGAAGCTATCA	AGGTAGCGGAAGTCTAATGTATATAGCAATACACTAAGA		600
QY	682	ATTTAACCGTAAACCCCTCT	CTCAACTATTAGTAAATGACGAGACATAA	CTTTTGGATAAAAAC	741
DB	601	ATTTAACCGTAAACCCCTCT	CTCAACTATTAGTAAATGACGAGACATAA	CTTTTGGATAAAAAC	660
QY	742	AAT	744		
DB	661	AAT	663		

RESULT 15  
AR193049 AR193049 663 bp DNA linear PAT 20-APR-2002  
LOCUS Sequence 1 from patent US 6346407.  
DEFINITION AR193049  
ACCESSION AR193049  
VERSION AR193049.1 GI:20239014  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 663)  
De Buyl, E., Lahaye, A., Ledoux, P. and Detroz, R.  
Xylanase, microorganisms producing it, DNA molecules, methods for  
preparing this xylanase and uses of the latter  
JOURNAL Patent: US 6346407-A 1 12-FEB-2002;  
FEATURES  
Location/Qualifiers  
source 1..663  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 89.1%; Score 663; DB 6; Length 663;  
Best Local Similarity 100.0%; Pred. No. 7.7e-166;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 82 CAATCGTCACGACAAATCCATTGGCAACCAACGATGGCTATGATTGAATTTGGAAA 141  
DB 1 CAATCGTCACGACAAATCCATTGGCAACCAACGATGGCTATGATTGAATTTGGAAA 60  
QY 142 GATACGGTGGCTCTGGGACAAATGATCTCAATCATGGCGGTACGTTCAAGTGCCTCAATGG 201  
DB 61 GATACGGTGGCTCTGGGACAAATGATCTCAATCATGGCGGTACGTTCAAGTGCCTCAATGG 120  
QY 202 AACAAATGTAACAAATATTTCCGTAAAGGTAAAAAATTCATGAAACCAACACAC 261  
DB 121 AACAAATGTAACAAATATTTCCGTAAAGGTAAAAAATTCATGAAACCAACACAC 180  
QY 262 CAACAAATGTAACAAATGTAACAAATGTAACAAATGTAACAAATGTAACAAATGTAAC 321  
DB 181 CAACAAATGTAACAAATGTAACAAATGTAACAAATGTAACAAATGTAACAAATGTAAC 240  
QY 322 TATTTATCGCTATGTTGGACTGTTGACCCCTCTTGTGCAATATTAATTTGTGACAGT 381  
DB 241 TATTTATCGCTATGTTGGACTGTTGACCCCTCTTGTGCAATATTAATTTGTGACAGT 300  
QY 382 TGGGCAACTGGCGTCCACAGGAGCAACGGCTAAGGGGACCATCCTGTTGATGGAGGA 441  
DB 301 TGGGCAACTGGCGTCCACAGGAGCAACGGCTAAGGGGACCATCCTGTTGATGGAGGA 360  
QY 442 ACATATGATATCTAGAGACTCTTAGAGTCAATCAACCCCTCATTAAAGGGATTGCCACA 501  
DB 361 ACATATGATATCTAGAGACTCTTAGAGTCAATCAACCCCTCATTAAAGGGATTGCCACA 420  
QY 502 TTAAACAAATATTTGGAGTGTTCGAAGATCGAAACCGACAGTGGCAGCATTTCTGTACG 561  
DB 421 TTAAACAAATATTTGGAGTGTTCGAAGATCGAAACCGACAGTGGCAGCATTTCTGTACG 480  
QY 562 AACCACTTTAGAGCTGGGAAACCTTAGGGATGAATATGGGAAATGTATGAAGTCGCG 621  
DB 481 AACCACTTTAGAGCTGGGAAACCTTAGGGATGAATATGGGAAATGTATGAAGTCGCG 540  
QY 622 CTTACTGTAGAGGCTATCAAGTAGCGGAGTGTCTATGTATATAGCAATACACTAAGA 681  
DB 541 CTTACTGTAGAGGCTATCAAGTAGCGGAGTGTCTATGTATATAGCAATACACTAAGA 600  
QY 682 ATTAACGGTAAACCTCTCTCAACTATTAGTAAATGACGAGCATTAACCTTTGGATAAAAC 741  
DB 601 ATTAACGGTAAACCTCTCTCAACTATTAGTAAATGACGAGCATTAACCTTTGGATAAAAC 660  
QY 742 AAT 744  
|||

Db 661 AAT 663

Search completed: November 10, 2004, 21:26:28  
Job time : 3600.04 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 15:55:17 ; Search time 471.947 Seconds  
(without alignment)  
8275.436 Million cell updates/sec

Title: US-09-909-207-4  
Perfect score: 744  
Sequence: 1 ATGAGCAAAAGAAATTGAC.....TAACCTTGGTAAACAAAT 744

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 23Sep04: \*  
1: geneseqn1980s: \*  
2: geneseqn1990s: \*  
3: geneseqn2000s: \*  
4: geneseqn2001as: \*  
5: geneseqn2001bs: \*  
6: geneseqn2002as: \*  
7: geneseqn2002bs: \*  
8: geneseqn2003as: \*  
9: geneseqn2003bs: \*  
10: geneseqn2003cs: \*  
11: geneseqn2003ds: \*  
12: geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	744	100.0	744	2	AAT16102
2	744	100.0	1513	2	AAT16103
3	713.6	95.9	744	2	AAT13067
4	713.6	95.9	871	2	AAT30255
5	671.6	90.3	744	2	AAQ92878
6	663	89.1	663	2	AAT16101
7	587.2	78.9	747	2	AAT13074
8	393	52.8	1068	12	ADJ34947
9	276.6	37.2	1956	12	ADJ35051
10	274.4	36.9	684	12	ADJ34969
11	273	36.7	1022	2	AAQ80923
12	255.6	34.4	2364	3	AAZ51821
13	212.2	28.5	1190	2	AAT90972
14	204.8	27.5	1244	2	AAT08142
15	177.8	23.9	747	12	ADJ35011
16	154.4	20.8	164	2	AAQ92876
17	151.2	20.3	164	2	AAQ92875
18	143.6	19.3	1695	12	ADJ35101
19	142.6	19.2	1338	12	ADJ35151
20	136.2	18.3	1077	12	ADJ34965
21	130.2	17.5	1065	12	ADJ34999

22	128.4	17.3	1047	12	ADJ34949	Adj34949 DNA encod
23	121.2	16.3	678	12	ADJ34955	Adj34955 DNA encod
24	119.4	16.0	1375	2	AAQ90405	AAQ90405 Actinomad
25	119.4	16.0	1375	2	AAT64930	AAT64930 Actinomad
26	119	16.0	229	2	AAV15063	AAV15063 Xylanase
27	118.4	15.9	573	2	AAV36098	AAV36098 DNA seque
28	117.8	15.8	1207	2	AAT42374	AAT42374 Bacillus
29	117.6	15.8	1041	12	ADJ34943	Adj34943 DNA encod
30	116.8	15.7	1047	12	ADJ35083	Adj35083 DNA encod
31	115.2	15.5	213	2	AAV15059	AAV15059 Xylanase
32	115.2	15.5	596	3	AAA48219	AAA48219 T. reesei
33	115.2	15.5	596	6	AAAD29410	AAAD29410 Trichoder
34	115.2	15.5	596	9	AAAL60925	AAAL60925 Trichoder
35	115.2	15.5	596	9	AAAL60941	AAAL60941 Trichoder
36	112	15.1	1074	12	ADJ34983	Adj34983 DNA encod
37	112	15.1	1137	12	ADJ35039	Adj35039 DNA encod
38	111.8	15.0	1273	2	AAQ90388	AAQ90388 Xylanase
39	111.4	15.0	636	12	ADJ34981	Adj34981 DNA encod
40	111.4	15.0	942	10	ABQ80366	ABQ80366 A. fumiga
41	111.2	14.9	1008	12	ADJ35015	Adj35015 DNA encod
42	110.8	14.9	669	12	ADJ34951	Adj34951 DNA encod
43	110.4	14.8	1047	12	ADJ35037	Adj35037 DNA encod
44	110	14.8	1002	10	ABQ80365	ABQ80365 A. fumiga
45	109.4	14.7	1041	12	ADJ34953	Adj34953 DNA encod

## ALIGNMENTS

RESULT 1  
AAT16102  
ID AAT16102 standard; DNA; 744 BP.

AC AAT16102;

DT 16-OCT-2003 (revised)  
DT 15-MAY-1996 (first entry)

DE Xylanase precursor gene.

XX Xylanase; thermostable enzyme; paper; pulp; bleaching; feedstuffs; baking;  
KW ss.

XX Bacillus sp; strain 710/1 (LMG P-14798).

XX Key Location/Qualifiers

FT sig\_peptide 1..81

FT mat\_peptide 82..744

FT /tag= a

FT /tag= b

FT /EC\_number= "3.2.1.8"

XX AU9525086-A.

XX 08-FEB-1996.

XX 19-JUL-1995; 95AU-00025086.

XX 26-JUL-1994; 94BB-00000706.

XX 17-MAY-1995; 95BE-00000448.

XX (SOLV ) SOLVAY SA.

XX De Buyl E, Lahaye A, Ledoux P, Detroz R;

XX WPI; 1996-117341/13.

XX P-PSDB; AAR92054.

XX Bacillus derived xylanase active over wide pH range - used in treatment

XX of paper pulp, animal feeds and in bakery goods.

XX Claim 9; Page 54-55; 94pp; English.



RESULT 2  
AAT16103  
ID AAT16103 standard: DNA: 1513 BP.

	Query Match	100.0%;	Score 744;	DB 2;	Length 1513;
	Best Local Similarity	100.0%;	Pred. No. 2.1e-210;		
	Matches 744;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGAGACAAAGAATAATTGACGTGGTATTGCTTTTAGTCCTTTTTAGTTTGTTGTTGCATAAACCTTA	60		
DB	620	ATGAGACAAAGAATAATTGACGTGGTATTGCTTTTAGTCCTTTTTAGTTTGTTGTTGCATAAACCTTA	679		
QY	61	CCTGCAGAAATAAATTCAGGCACAATAATCGTCACCGACAATTCCAATTGGCAACCACGATGCC	120		
DB	680	CCTGCAGAAATAAATTCAGGCACAATAATCGTCACCGACAATTCCAATTGGCAACCACGATGCC	739		
QY	121	TATGATTATGAATTTTGGAAAAGATAGCGGTGGCTCTGGGCAATGATCTCAATCATGGC	180		
DB	740	TATGATTATGAATTTTGGAAAAGATAGCGGTGGCTCTGGGCAATGATCTCAATCATGGC	799		

RESULT 2  
AAT16103  
ID AAT1



Qy	181	GGTACGTTTCAGTCCCAATCGGAACAATGTTTAA	240
Db	800	GGTACGTTTCAGTCCCAATCGGAACAATGTTTAA	859
	241	TTCAATGAAACAAACACACACCAACAAAGTTG	300
Db	860	TTCAATGAAACAAACACACCAACAAAGTTG	919
	301	TTCCAAACCAATGGTAAATGCGTATTTATGCGT	360
Db	920	TTCCAAACCAATGGTAAATGCGTATTTATGCGT	979
	361	GAATATATATTTGTGACAGTTGGGGCAACTGG	420
Db	980	GAATATATATTTGTGACAGTTGGGGCAACTGG	1039
	421	ACCATCACTCTTGATGGAGGAAACATATGATA	480
Db	1040	ACCATCACTCTTGATGGAGGAAACATATGATA	1099
	481	TCCATTAAAGGGGATTCGCACATTTAAACAAT	540
Db	1100	TCCATTAAAGGGGATTCGCACATTTAAACAAT	1159
	541	AGTGGCAGATTTCTGTACGAAACCACTTTAG	600
Db	1160	AGTGGCAGATTTCTGTACGAAACCACTTTAG	1219
	601	GGGAAATGTATGAAAGTCGCGCTTACTGTAG	660
Db	1220	GGGAAATGTATGAAAGTCGCGCTTACTGTAG	1279
	661	GTATATAGCAATACACTAAGAAATTAACCGGT	720
Db	1280	GTATATAGCAATACACTAAGAAATTAACCGGT	1339
	721	AGCATATCTTTGGATAAAAAACAAT	744
Db	1340	AGCATATCTTTGGATAAAAAACAAT	1363

RESULT 3	
AAV13067	
ID	AAV13067 standard; DNA; 744 BP.
XX	
XX	AAV13067;
XX	
DT	19-MAY-1998 (first entry)
XX	
DE	Glycosyl hydrolase family 11 xylanase DNA derived from <i>Bacillus</i> sp.
XX	
KW	<i>Bacillus</i> sp; xylanase; glycosyl hydrolase family 11; isolation;
KW	microorganism; identification; ss.
XX	
OS	<i>Bacillus</i> sp.

	Key	Location/Qualifiers
XA	CDS	1..747
FH		/tag= a
FT		/product= "glycosyl hydrolase family 11 xylanase"
FT		
FT		
XX		
XX		
PN	WO9743409-A2.	
XX		
XX		
PD	20-NOV-1997.	
XX		
PF	12-MAY-1997;	97WO-DK000216.
XX		
PR	10-MAY-1996;	96DK-00000562.
XX		
PA	(NOVO ) NOVO-NORDISK AS.	
XX		
PI	Dalbøge H, Diderichsen B, Sandal T, Kauppinen S;	
XX		

**WPI; 1998-008878/01.**  
**P-PSDB; AAW44262.**  
 Isolating novel DNA sequences from microorganisms - without the need for culturing the microorganism.  
 Example 1; Page 31-32; 72pp; English.  
 The present sequence encodes a polypeptide with xylanase activity used in an example of the present invention. The present invention describes a novel method for providing a novel DNA sequence encoding a polypeptide from a microorganism with an activity of interest. The method comprises: (i) PCR amplification of the DNA with PCR primers with homology to (a) known gene(s) encoding a polypeptide with an activity of interest; (ii) linking the obtained PCR product of a 5' structural gene sequence and a 3' structural gene sequence; (iii) expressing the resulting hybrid DNA sequence; (iv) screening for hybrid DNA sequences encoding a polypeptide with the activity of interest or a related activity; and (v) isolating the hybrid DNA sequence identified in step (iv). This method provides for the identification and isolation of sequences from microorganisms without having to cultivate and isolate the microorganism.  
 Sequence 744 BP; 249 A; 146 C; 157 G; 192 T; 0 U; 0 Other;

Query Match	95.9%	Score 713.6	DB 2	Length 744
Best Local Similarity	97.4%	Pred. No. 1.7e-201		
Matches 725	Conservative 0	Mismatches 19	Indels 0	Gaps 0
QY 1	ATGAGACAAAGAAATTTGACGTGATTTTACGCTTTTGTAGTCTTTTGTGTTGTTGACATAACCTTA	60		
DB 1				
QY 1	ATGAGACAAAGAAATTTGACGTGATTTTACGCTTTTGTAGTCTTTTGTGTTGTTGACATAACCTTA	60		
DB 1				
QY 61	CCTGCAGAAATATTTACGGACACAATCGTCACCGACAATTCATTTGCCAACCCACGATGGC	120		
DB 61	CCTGCAGAAATATTTACGGACACAATCGTCACCGACAATTCATTTGCCAACCCACGATGGC	120		
QY 121	TATGATTTATGAATTTTGGAAAGATAGCGGTGGCTCTGGGACAATGATTTCTCAATCATGGC	180		
DB 121	TATGATTTATGAATTTTGGAAAGATAGCGGTGGCTCTGGGACAATGATTTCTCAATCATGGC	180		
QY 181	GGTACGTTTCAGTGCCTTAAAGCAATGTTTAAACAATATTTATTCCTGAAGGTAAAAA	240		
DB 181	GGTACGTTTCAGTGCCTTAAAGCAATGTTTAAACAATATTTATTCCTGAAGGTAAAAA	240		
QY 241	TTCAATGAACACAAACACACCAACAAGTTGGTAACATGTCCATAAACTACGGAGCCAAAC	300		
DB 241	TTCAATGAACACAAACACACCAACAAGTTGGTAACATGTCCATAAACTACGGAGCCAAAC	300		
QY 301	TTCCAAACAAATGTAATGCGTATTTATGCGTCTATGTTGGACTGTGACCTCTTGTCT	360		
DB 301	TTCCAGGCAAAACGAAATGCGTATTTATGCGTCTATGTTGGACTGTGACCTCTTGTCT	360		
QY 361	GAATTTATTTATTTGTCGACAGTTGGGGCAACTGGCGTCCACAGAGAGCAACGCTTAAGGG	420		
DB 361	GAATTTATTTATTTGTCGACAGTTGGGGCAACTGGCGTCCACAGAGAGCAACGCTTAAGGG	420		
QY 421	ACCATCACTGTTGATGGAGGAACAATATGATATCTTACGAGACTCTTAGAGTCAATCAACCC	480		
DB 421	ACCATCACTGTTGATGGAGGAACAATATGATATCTTACGAGACTCTTAGAGTCAATCAACCC	480		
QY 481	TCCATTAAAGGGGATGCGCAATTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACG	540		
DB 481	TCCATTAAAGGGGATGCGCAATTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACG	540		
QY 541	AGTGGCAGGATTTCTGTCAGCAACCACTTTAGCGGTGGGAAAACCTTAGGGATGAAATG	600		
DB 541	AGTGGCAGCAATTTCTGTCAGCAACCACTTTAGCGGTGGGAAAACCTTAGGGATGAAATG	600		
QY 601	GGGAAAATGTATGAAGTCGCGCTTACTGTAGAGGCTATCAAACTAGCGGAAGTGCTAAT	660		
DB 601	GGGAAAATGTATGAAGTCGCGCTTACTGTAGAGGCTATCAAACTAGCGGAAGTGCTAAT	660		
QY 661	GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTTAGTATGACGAG	720		

[illegible]

PD 06-JUL-1995. |||||TCCATTAGGGGATTCACCAATTTAAACAATATTGGAGTGTCCGAAGATCGAAACGCACG 537  
 XX 23-DEC-1994; 94WO-BP004312. ||||| 541 AGTGCCAGATTTCTGTCTAGCAACACCTTTAGAGCGTGGGAAACTTAGGGATGAATATG 600  
 XX 24-DEC-1993; 93EP-00203694. ||||| 538 AGTGCCACAAATTTCTGTCTAGCAACACCTTTAGAGCGTGGGAAACTTAGGGATGAATATG 597  
 XX (KONN ) GIST-BROCADES NV. ||||| 601 GGGAAATGTATGAAGTCGCGCTTACTGTAGAGGCTATCAAGTAGCGGAAAGTGTCTAAT 660  
 XX Van Solingen P, Williams DP, Iverson S, Farrell RL, Herbes WT; ||||| 598 GGGAAATGTATGAAGTCGCGCTTACTGTAGAGGCTATCAAGTAGCGGAAAGTGTCTAAT 657  
 XX Goedegebuur F, Jones BE; ||||| 661 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCACTATTAGTATGACGAG 720  
 XX WPI: 1995-246385/32. ||||| 658 GTATATAGCAATACACTAAGAAATTAACGGGAAACCTCTCTCACTATTAGTATGACGAG 717  
 XX P-PSDB; AAR76551. ||||| 721 AGCATAACTTGGTATTAACCAAT 744  
 XX Novel xylanase enzyme active at high pH - useful in paper and pulp prodn. ||||| 718 AGCATAACTTAGATAAAACAAT 741  
 XX processes. |||||  
 XX Claim 4; Page 42-43; 54pp; English.  
 XX The sequence encodes an alkaline endo-1,4-beta-D-xylanase (G-type) from  
 CC Bacillus sp. 1-43-3. Fragments of the sequence (e.g. the internal  
 CC fragments given in AAQ92876 and AAQ92877) may be amplified by polymerase  
 CC chain reaction, e.g. using primers with sequences AAQ92866, AAQ92867,  
 CC AAQ92868 and AAQ92869. The DNA may be cloned in Escherichia coli using a  
 CC plasmid vector for recombinant xylanase production. The xylanase may be  
 CC used in the paper and pulp industries, where it produces an increase in  
 CC ISO brightness of softwood pulp of at least 1.0 over non-enzymatically  
 CC treated pulp in an ECF pulp bleaching process, at pH 9.0 and 65 deg C.  
 CC The enzyme may be used in production of paper, board and fluff pulp, and  
 CC has low cellulase activity. The increased brightness produced using the  
 CC xylanase allows reduction in the amount of bleaching chemicals used.  
 CC (Updated on 16-OCT-2003 to standardise OS field)  
 XX Sequence 744 BP; 248 A; 145 C; 159 G; 192 T; 0 U; 0 Other;  
 SQ  
 Query Match 90.3%; Score 671.6; DB 2; Length 744;  
 Best Local Similarity 95.7%; Pred. No. 5.1e-189;  
 Matches 712; Conservative 0; Mismatches 29; Indels 3; Gaps 2;  
 QY 1 ATGAGACAAAGAAATGACGTTGATTTTATGAGCTTTTATGTTTGTGTTTGCACCTTA 60  
 DB 1 ATGAGCCAAAGAAATGACGTTGA--TTAACCTTTTATGTTG--TTTGCACCTTAACCTTA 57  
 QY 61 CTGCGAGAATAATTCAGGCACAAATCGTCACGCAATTCGATTCGCAACACGATGCG 120  
 DB 58 CTGCGAAGAATAATGTCAGGCACAAATCGTCACGCAATTCGATTCGCAACACGATGCG 117  
 QY 121 TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTGGGACAAATGATTTCTCAATCATGGC 180  
 DB 118 TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTGGGACAAATGATTTCTCAATCATGGC 177  
 QY 181 GGTACGTTTCAGTCCCAATGGAACAATGTTTAAACAATATATTTCGTAAGGTAAAAA 240  
 DB 178 GGTACGTTTCAGTCCCAATGGAATAATGTTTAAACAATATATTTCGTAAGGTAAAAA 237  
 QY 241 TCAATGAACAACAACAACAACAAGTTGGTAAACATGTCATTAACACTACGAGCCAC 300  
 DB 238 TCAATGAACAACAACAACAACAAGTTGGTAAACATGTCATTAACACTACGAGCCAC 297  
 QY 301 TTCACCAATGATGATGCTATTTATGCTCTATGTTGGACTGTGTAACCTCTTGTG 360  
 DB 298 TTCACCAATGATGATGCTATTTATGCTCTATGTTGGACTGTGTAACCTCTTGTG 357  
 QY 361 GAATATTATATTGTCAGTGTGGGCAACTGGCGTCCACAGAGCAACGCTTAAGGGG 420  
 DB 358 GAATATTATATTGTCAGTGTGGGCAACTGGCGTCCACAGAGCAACGCTTAAGGGG 417  
 QY 421 ACCATCACTGTTGATGAGGAAACATATGATATCTACAGACTCTTAGATCAATCAACC 480  
 DB 418 ACCATCACTGTTGATGAGGAAACATATGATATCTTAGAACTCTTAGATCAATCAACC 477  
 QY 481 TCCATTAAAGGGGATGCGACATTTAAACAATATTGGAGTGTTCGAAGATCGAAACGCACG 540

RESULT 6

AAT16101

ID AAT16101 standard; DNA; 663 BP.

XX

AC AAT16101;

XX

DT 16-OCT-2003 (revised)

DT 15-MAY-1996 (first entry)

XX

DE Xylanase gene.

XX

KW Xylanase; thermostable enzyme; paper; pulp; bleaching; feedstuff; baking;

KW ss.

XX

OS Bacillus sp; strain 710/1 (LMG P-14798).

XX

FH Key Location/Qualifiers

FT mat\_peptide 1..663

FT /tag= a

FT /EC\_number= "3.2.1.8"

XX

PN AU9525086-A.

XX

PD 08-FEB-1996.

XX

PF 19-JUL-1995; 95AU-00025086.

XX

PR 26-JUL-1994; 94BB-00000706.

XX

PR 17-MAY-1995; 95BE-00000448.

XX

PA (SOLV ) SOLVAY SA.

XX

PI De Buyl E, Lahaye A, Ledoux P, Detroz R;

XX

XX WPI; 1996-117341/13.

DR

DR P-PSDB; AAR92053.

XX

XX Bacillus derived xylanase active over wide pH range - used in treatment

PT

PT of paper pulp, animal feeds and in bakery goods.

XX

PS Claim 30; Page 50-51; 94pp; English.

XX

CC A DNA sequence (AAT16101) coding for a thermostable mature xylanase  
 CC (AAR92053) was isolated from a gene library of Bacillus sp. 720/1 (LMG P-  
 CC 14798). Sequences were also obtd. (see AAT16102 and AAT16103) for the  
 CC xylanase precursor and for the complete gene including 5' and 3'  
 CC untranslated sequences. DNA coding for the mature enzyme may be  
 CC incorporated into a vector and expressed from either its own promoter or  
 CC from the Bacillus pumilus PRL B12 promoter (AAQ73996), and used for  
 CC prodn. of recombinant xylanase in transformed hosts, pref. Bacillus  
 CC licheniformis or B. pumilus. The enzyme is useful in the paper-pulp,  
 CC animal feed and baking industries. (Updated on 16-OCT-2003 to standardise  
 CC OS field)

Fri Nov 12 14:59:40 2004

us-09-909-207-4.rng

```

XX SQ Sequence 663 BP; 220 A; 131 C; 146 G; 166 T; 0 U; 0 Other;
Query Match      89.1%; Score 663; DB 2; Length 663;
Best Local Similarity 100.0%; Pred. NO. 1.7e-186;
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 82 CAATTCGTCACCGCAATTCATTTGGCAACCAACGATGGCTATGATTAATGAAATTTGGAAA 141
DB 1 CAATTCGTCACCGCAATTCATTTGGCAACCAACGATGGCTATGATTAATGAAATTTGGAAA 60
QY 142 GATAGCGTGGCTCTGGGACAAATGATTTCAATCATGGCGGTACGTTCACTGCCAATGG 201
DB 61 GATAGCGTGGCTCTGGGACAAATGATTTCAATCATGGCGGTACGTTCACTGCCAATGG 120
QY 202 AACATGTTAACACATATTTCCGTAAGGTAAATAATTCATGAAACACAAACACAC 261
DB 121 AACATGTTAACACATATTTCCGTAAGGTAAATAATTCATGAAACACAAACACAC 180
QY 262 CAACAAGTTGGTAACATGTCATATAAATACGAGGCAACATTTCCAAACCAATGGTAATGCG 321
DB 181 CAACAAGTTGGTAACATGTCATATAAATACGAGGCAACATTTCCAAACCAATGGTAATGCG 240
QY 322 TATTTATGGCTATGGTTGACCTTTGACCTCTTGTGCAATATTAATTTGTCGACAGT 381
DB 241 TATTTATGGCTATGGTTGACCTTTGACCTCTTGTGCAATATTAATTTGTCGACAGT 300
QY 382 TGGGCAACTGGCGTCCACGAGGCAACCGCTAAGGGGACCATCTGTTGATGGAGGA 441
DB 301 TGGGCAACTGGCGTCCACGAGGCAACCGCTAAGGGGACCATCTGTTGATGGAGGA 360
QY 442 ACATATGATATACGAGACTTTAGAGTCAATCAACCTCCATTAAGGGGATGGCCACA 501
DB 361 ACATATGATATACGAGACTTTAGAGTCAATCAACCTCCATTAAGGGGATGGCCACA 420
QY 502 TTAAACATATTTGGAGTCTCGAAGATCGAAACCGACGAGTGGCAGCATTTCTGACG 561
DB 421 TTAAACATATTTGGAGTCTCGAAGATCGAAACCGACGAGTGGCAGCATTTCTGACG 480
QY 562 AACCACTTTAGAGCGTGGGAAACCTTAGGATGAATATATGGGAAATATGTAAGTCCGG 621
DB 481 AACCACTTTAGAGCGTGGGAAACCTTAGGATGAATATATGGGAAATATGTAAGTCCGG 540
QY 622 CTATCTGTAGAGGCTATCAAGTATGAGCGAAGTCTAATGATATAGCAATACACTAAGA 681
DB 541 CTATCTGTAGAGGCTATCAAGTATGAGCGAAGTCTAATGATATAGCAATACACTAAGA 600
QY 682 ATTAACGGTAACCTCTCTCAACTATTAGTATGACGAGACATAACTTTGGATATAAAC 741
DB 601 ATTAACGGTAACCTCTCTCAACTATTAGTATGACGAGACATAACTTTGGATATAAAC 660
QY 742 AAT 744
DB 661 AAT 663

```

```

RESULT 7
AAV13074
ID AAV13074 standard; DNA; 747 BP.
AC AAV13074;
XX AAV13074;
DT 19-MAY-1998 (first entry)
XX Xylanase activity positive clone DNA SEQ ID NO:11.
DE Bacillus sp; xylanase; glycosyl hydrolase family 11; isolation;
KW microorganism; identification; hybrid DNA; ss.
XX Synthetic.
OS Bacillus sp.
XX Key Location/Qualifiers
FH

```

```

FT CDS
FT 1..747
FT /*tag= a
FT /product= "positive clone for xylanase activity"
XX WO9743409-A2.
XX 20-NOV-1997.
XX 12-MAY-1997; 97WO-DK000216.
XX 10-MAY-1996; 96DK-00000562.
XX (NOVO ) NOVO-NORDISK AS.
XX Dalboge H, Diderichsen B, Sandal T, Kauppinen S;
XX WPI; 1998-008878/01.
XX Isolating novel DNA sequences from microorganisms - without the need for
XX culturing the microorganism.
XX Example 1; Page 35; 72pp; English.
XX The present sequence represents a positive clone for xylanase activity
XX from an example of the present invention. The present invention describes
XX a novel method for providing a novel DNA sequence encoding a polypeptide
XX from a microorganism with an activity of interest. The method comprises:
XX (i) PCR amplification of the DNA with PCR primers with homology to (a)
XX known gene(s) encoding a polypeptide with an activity of interest; (ii)
XX linking the obtained PCR product of a 5' structural gene sequence and a
XX 3' structural gene sequence; (iii) expressing the resulting hybrid DNA
XX sequence; (iv) screening for hybrid DNA sequences encoding a polypeptide
XX with the activity of interest or a related activity; and (v) isolating
XX the hybrid DNA sequence identified in step (iv). This method provides for
XX identification and isolation of sequences from microorganisms without
XX having to cultivate and isolate the microorganism
XX Sequence 747 BP; 236 A; 168 C; 154 G; 189 T; 0 U; 0 Other;
Query Match      78.9%; Score 587.2; DB 2; Length 747;
Best Local Similarity 86.8%; Pred. NO. 6.3e-164;
Matches 646; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 1 ATGAGACAAAAGAAATTTGAGTTCATTTTACCTTTTGTGTTTGTGCACTAACCTTA 60
DB 1 ATGAGACAAAAGAAATTTGAGTTCATTTTACCTTTTGTGTTTGTGCACTAACCTTA 60
QY 61 CCTCAGAAATATTTGAGGCAAAATCGTCAACGACAAATTCATTTGGCAACCAACGATGGC 120
DB 61 CCTCAGAAATATTTGAGGCAAAATCGTCAACGACAAATTCATTTGGCAACCAACGATGGC 120
QY 121 TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTGGGACAAATGATTTCAATCATGGC 180
DB 121 TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTGGGACAAATGATTTCAATCATGGC 180
QY 181 GGTACGTTCACTGCCCAATGGAACAAATTTTAAACACATATTTTCCGTAAGGTAAATAA 240
DB 181 GGTACGTTCACTGCCCAATGGAACAAATTTTAAACACATATTTTCCGTAAGGTAAATAA 240
QY 241 TTCAATGAAACACAAACACACCAACCAAGTGTGTAACATGTCCATAAACTACGGAGCCAAC 300
DB 241 TTCAATGAAACACAAACACACCAACCAAGTGTGTAACATGTCCATAAACTACGGAGCCAAC 300
QY 301 TTCCAAACCAATGGTAATGGCTATTTATGCGTCTATGGTGGACTGTTGACCTCTTCTG 360
DB 301 TTCCAAACCAATGGTAATGGCTATTTATGCGTCTATGGTGGACTGTTGACCTCTTCTG 360
QY 361 GAATATTATTTGTCGACAGTGTGGGCAACTGGCGTCCACGAGGACCAACGCTTAAGGG 420
DB 361 GAATATTATTTGTCGACAGTGTGGGCAACTGGCGTCCACGAGGACCAACGCTTAAGGG 420
QY 421 ACCATCACTGTTGATGGAGGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACC 480

```

Db 421 ACCATCACTGTTGACGGGGGAGGTATGATATCTTACAAGCAACCAAGGTCAATCAGCCA 480  
QY 481 TCATTAAGGGGATGCGACATTTAAACATATTCGAGTGTTCGAAGATCGAAACGACG 540  
Db 481 TCTATTCAGGGCAGCGCCACCTTCAATCAGTACTGTCGATTCGACAGACGACGCGACC 540  
QY 541 AGTGGCAGATTTCTGTCAGCAACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATG 600  
Db 541 AGCGGCACGTGCTACACGCGCAACCACTTTAATGCTGGCTGCTCTGGCATGATATG 600  
QY 601 GGGAAATGTATGAAGTGCCTTACTGTAGAGGCTATCAAGTAGCGGAAGTCTAAT 660  
Db 601 GGTGCAATCAATACCAAGTCTGTTACTGAGGGCTACCAATCTACCGAAGTCTAAT 660  
QY 661 GTATATAGCAATACACTAAGATTAACGGTAAACCTCTCTCAACTATTAGTAATGACGAG 720  
Db 661 GTATATAGCAATACACTAAGATTAACGGTAAACCTCTCTCAACTATTAGTAATGACGAG 720  
QY 721 AGCATAACTTTGCGTAAACCAAT 744  
Db 721 AGCATAACTCTAGATAAACCAAT 744

## RESULT 8

ADJ34947  
ID ADJ34947 standard; DNA; 1068 BP.

XX AC ADJ34947;

XX 22-APR-2004 (first entry)

XX DNA encoding xylanase from an environmental sample seq id 163.

XX antibacterial; fungicide; thermostable xylanase activity;  
KW dough conditioning; beverage production; nutritional supplement;  
KW animal feed; lignin reduction; wood product; xylan; bacterial infection;  
KW fungal infection; coccidiosis; gene; ds.

XX Unidentified.

XX WO2003106654-A2.

XX 24-DEC-2003.

XX 16-JUN-2003; 2003WO-US019153.

XX 14-JUN-2002; 2002US-0389299P.

XX (DIVE-) DIVERSA CORP.

XX Steer B, Callen W, Healey S, Hazlewood G, Wu D, Blum D;

XX Esteghlalian A;

XX WPI; 2004-099016/10.

XX P-PSDB; ADJ34948.

XX Novel xylanase recombinant polypeptide useful for improving textile

XX texture, treating paper, eliminating microorganisms.

XX Claim 1; SEQ ID NO 163; 570pp; English.

XX The invention describes an isolated or recombinant polypeptide (I),  
CC having 50% or more identity to 190 300-1200 residue amino acid sequences  
CC (S1), given in the specification, over a region of 100 or more residues  
CC and the polypeptide as thermostable xylanase activity. (I) is useful for:  
CC dough conditioning; beverage production; as a nutritional supplement in  
CC animal feed; reducing lignin in a wood or a wood product; and for  
CC eliminating and protecting animals from a microorganism comprising xylan.  
CC The polynucleotide (II) encoding (I) is useful for amplifying nucleic  
CC acid encoding a polypeptide having a xylanase activity which involves  
CC amplification of a template nucleic acid with a primer pair capable of  
CC amplifying (II) or its subsequence. (I) is useful for treating and  
CC preventing bacterial infection and fungal infection e.g. coccidiosis.

CC This sequence encodes xylanase protein isolated from an environmental  
CC sample.

XX Sequence 1068 BP; 303 A; 249 C; 279 G; 237 T; 0 U; 0 Other;

SQ Query Match 52.8%; Score 393; DB 12; Length 1068;

Best Local Similarity 75.2%; Pred. No. 3.9e-106;

Matches 518; Conservative 0; Mismatches 165; Indels 6; Gaps 2;

QY 1 ATGACACAAAAGAAATGACGCTGATTTTACGCTTTTGTAGTTTGTGCTACCTTA 60

Db 1 ATGAAGCAAAAGAAATGAGTTGTTGCCGCAITTTTCTCTGTTTACGTTGCATTT 60

QY 61 CCTGACAGAAATTAATTCAGGCAAAATCGTCACCGCAAAATTCATTTGGCAACCAACGATGC 120

Db 61 CCTGGG---CAGTGCATGCGCAGACGATCCAGCAATTCGGTCGGTACGATACGCT 117

QY 121 TATGATTAATGATTTTGGAAAGATAGCGGTGGCTCTGGGACAAATGATTTCTCAATCATGCG 180

Db 118 TATGACTATGAATACTGGAAGGACAGC---GGGAATGGAACATATGTTTCTCGGTAGTGGC 174

QY 181 GGTAGTTTCAGTCCCAATGGAACAAATGTTAAACACATATATTCCGTAAAGGTAAAAA 240

Db 175 GGTAGTTTCAGTCCCAATGGAACAAATGTTAAACACATATATTCCGTAAAGGTAAAAA 234

QY 241 TTCAATGAAACACAAACACACCAACAAAGTTGGTAAACATGTCCATAAACTACGAGCCCAAC 300

Db 235 TTCAATGAGACGCGACGCCATCAGCAATTTGGAACATTTCCATAACCTTATGTCGACCC 294

QY 301 TTCCAACCAAAATGGTAAATGGGTATTTATGCGTCTATGGTTGACCTGTTGACCTCTTGTCT 360

Db 295 TACCAACCGAATGGCAATTCGTTATTAACGGTCTATGGCTGGACGGTTGACCCCTCGTC 354

QY 361 GAATATTATTTGTCGACAGTTGGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGGG 420

Db 355 GAATATTATTTGTCGACAGTTGGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGGG 414

QY 421 ACCATCACTGTTGATGGAGGAACATATGATATCTACGAGACTCTTTAGAGTCAATCAACCC 480

Db 415 ACTGTTAAGCTTGACGAGGAACGATGACATTTATGAGCAACTCGTCTCAACGACCT 474

QY 481 TCCATTAAGGGGATTTGCCACATTTAAACATATTTGAGTGTTCGAAGATCGAAACGACG 540

Db 475 TCCATTAAGGGGATTTGCCACATTTAAACATATTTGAGTGTTCGAAGATCGAAACGACG 534

QY 541 AGTGGCAGATTTCTGTCAGCAACCACTTTAGAGCGTGGGAAACTTAGCGATGAATATG 600

Db 535 AGCGGAACCATATCTGTAAGCGGACACTTTAAGGCTGGGGAATTTGGGATGACCATG 594

QY 601 GGGAAATGTATGAAGTCCGCTTACTGTAGAAGGCTATCAAAGTAGCGGAAGTGTCTAAT 660

Db 595 GCGAGATGTATGAAGTCCGCTTACGGTTGAAGGCTATCAAAGCAGTGAAGCGCTAAT 654

QY 661 GTATATAGCAATACACTAAGATTAACCG 689

Db 655 GTGATAGCCATACACTGACGATCGGCG 683

## RESULT 9

ADJ35051

ID ADJ35051 standard; DNA; 1956 BP.

XX AC ADJ35051;

XX 22-APR-2004 (first entry)

XX DNA encoding xylanase from an environmental sample seq id 267.

XX antibacterial; fungicide; thermostable xylanase activity;  
KW dough conditioning; beverage production; nutritional supplement;  
KW animal feed; lignin reduction; wood product; xylan; bacterial infection;  
KW fungal infection; coccidiosis; gene; ds.

OS Unidentified.  
XX WO2003106654-A2.  
XX 24-DEC-2003.  
XX  
XX 16-JUN-2003; 2003WO-US019153.  
XX 14-JUN-2002; 2002US-0389299P.  
XX (DIVE-) DIVERSA CORP.  
XX Steer B, Callen W, Healey S, Hazlewood G, Wu D, Blum D;  
PI Esteghlalian A;  
XX  
XX WPI; 2004-099016/10.  
XX P-PSDB; ADJ35052.  
XX Novel xylanase recombinant polypeptide useful for improving textile  
PT texture, treating paper, eliminating microorganisms.  
XX  
XX Claim 1; SEQ ID NO 267; 570pp; English.  
XX The invention describes an isolated or recombinant polypeptide (I),  
CC having 50% or more identity to 190 300-1200 residue amino acid sequences  
CC (SI), given in the specification, over a region of 100 or more residues  
CC and the polypeptide as thermostable xylanase activity. (I) is useful for:  
CC dough conditioning; beverage production; as a nutritional supplement in  
CC animal feed; reducing lignin in a wood or a wood product; and for  
CC eliminating and protecting animals from a microorganism comprising xylan.  
CC The polynucleotide (II) encoding (I) is useful for amplifying nucleic  
CC acid encoding a polypeptide having a xylanase activity which involves  
CC amplification of a template nucleic acid with a primer pair capable of  
CC amplifying (II) or its subsequence. (I) is useful for treating and  
CC preventing bacterial infection and fungal infection e.g. coccidiosis.  
CC This sequence encodes xylanase protein isolated from an environmental  
CC sample.  
XX  
XX Sequence 1956 BP; 614 A; 384 C; 423 G; 535 T; 0 U; 0 Other;  
SQ  
Query Match 37.2%; Score 276.6; DB 12; Length 1956;  
Best Local Similarity 66.2%; Pred. No. 2e-71;  
Matches 415; Conservative 0; Mismatches 209; Indels 3; Gaps 1;  
QY 69 AATAATTCAGGCACCAATCGTCACCGACAATTCCTATGGCAACCAACGATGGCTATGATTA 128  
DB 81 AGTACTCGCGGGCGAATTAATTACGACAAATGAGACAGGCACACATGGAGGCTACGACTA 140  
QY 129 TGAATTTTGGAAAGATAGCGGTGGCTCTGGGACAATGATTCATATCATGGCGGTACGTT 188  
DB 141 TGAGCTCTGMAAGA---CTACGGAAATACGATTATGGAACCTTAACGACGGTGTACTTT 197  
QY 189 CAGTGGCCAAATGGAACAATGTTAAACAATATATTCCTGTAAGGTAAATAAATTCGAATGA 248  
DB 198 TAGTTGTCATGAGTAATATCGGTAAATGACATATTTAGAAAGGGAGAAATTAATTC 257  
QY 249 AACACAAACACACCAAGTTGGTAACATGTCCATAAATACGAGCAACCTTCAAC 308  
DB 258 CGCAAAACCTATCAGAAATAGGAGATATAGTAGTTGATATGGCTGTGATTACAATCC 317  
QY 309 AATAGTGAATGCTATTTATGCTCTATGGTTGAGCTGTGACCTCTTGTGTAATTA 368  
DB 318 AAACGGAAATTCCTATTTGTGTTTACGGTTTGAACAAGAAATCCACTGGTTGAATTA 377  
QY 369 TATTGTCAGATGGGCAACTGGCGTCCACCAGAGCAACGCTTAAGGGACCAATCAC 428  
DB 378 CATTTAGAAAGCTGGGGAGCTGGCGTCCACTGGAGCAACACCCAAAGGACCAATCAC 437  
QY 429 TGTGATGGAGGAACATATGATATCTACGAGACTTTAGAGTCAATCAACCCCTCCATTA 488  
DB 438 AGTGATGGCGGTATTATGAATATATGAACACTACCCGGGTAAATCAGCCCTTCCATCGA 497  
QY 489 GGGGATTGCCACATTTTAAACAATATTGGAGTGTTCGAGATTCGAAGATCGAAACGACGAGTGGCAC 548

DB 498 TGGAACTGCGACATTCACAAATATTGGAGTGTTCGTACATCCAAAGAGAACAAAGCGAAC 557  
QY 549 GATTTCTGTGACCAACCACTTTAGGCGTGGGAAACTTAGGATGAATATGGGGAAAAAT 608  
DB 558 AATATCTGTCACTGAACATTTTAAACAGTGGGAAGAATGGGCAATGGGATGGGTAAGAT 617  
QY 609 GTATGAAGTCGGCTTTACTGTAGAGGCTATCAAGTAGCGGAAGTGTCTAATGTATATAG 668  
DB 618 GTATGAAGTTCCTTTACCGTTGAAGGTTATCAGAGCAGTGGGTACGCTAATGTATATA 677  
QY 669 CAATACACTAAGAAATTAACGGTAACCC 695  
DB 678 GAATGAATCAGAATAGGTGCAATCC 704  
RESULT 10  
ADJ34969  
ID ADJ34969 standard; DNA; 684 BP.  
XX AC ADJ34969;  
XX DT 22-APR-2004 (first entry)  
XX DE DNA encoding xylanase from an environmental sample seq id 185.  
XX KW antibacterial; fungicide; thermostable xylanase activity;  
XX KW dough conditioning; beverage production; nutritional supplement;  
KW animal feed; lignin reduction; wood product; xylan; bacterial infection;  
KW fungal infection; coccidiosis; gene; ds.  
XX OS Unidentified.  
XX PN WO2003106654-A2.  
XX PD 24-DEC-2003.  
XX PF 16-JUN-2003; 2003WO-US019153.  
XX PR 14-JUN-2002; 2002US-0389299P.  
XX PS (DIVE-) DIVERSA CORP.  
XX Steer B, Callen W, Healey S, Hazlewood G, Wu D, Blum D;  
PI Esteghlalian A;  
XX WPI; 2004-099016/10.  
XX P-PSDB; ADJ34970.  
XX Novel xylanase recombinant polypeptide useful for improving textile  
PT texture, treating paper, eliminating microorganisms.  
XX  
XX Claim 1; SEQ ID NO 185; 570pp; English.  
XX The invention describes an isolated or recombinant polypeptide (I),  
CC having 50% or more identity to 190 300-1200 residue amino acid sequences  
CC (SI), given in the specification, over a region of 100 or more residues  
CC and the polypeptide as thermostable xylanase activity. (I) is useful for:  
CC dough conditioning; beverage production; as a nutritional supplement in  
CC animal feed; reducing lignin in a wood or a wood product; and for  
CC eliminating and protecting animals from a microorganism comprising xylan.  
CC The polynucleotide (II) encoding (I) is useful for amplifying nucleic  
CC acid encoding a polypeptide having a xylanase activity which involves  
CC amplification of a template nucleic acid with a primer pair capable of  
CC amplifying (II) or its subsequence. (I) is useful for treating and  
CC preventing bacterial infection and fungal infection e.g. coccidiosis.  
CC This sequence encodes xylanase protein isolated from an environmental  
CC sample.  
XX  
XX Sequence 684 BP; 221 A; 129 C; 158 G; 176 T; 0 U; 0 Other;  
SQ  
Query Match 36.9%; Score 274.4; DB 12; Length 684;  
Best Local Similarity 64.6%; Pred. No. 5.8e-71;

Matches 442; Conservative 0; Mismatches 236; Indels 6; Gaps 2;

QY 1 ATGAGACAAAAGAAATTGAGTGGTGTATTTAGCTTTTATGTTTGTGCTACCTTA 60  
 DB 1 ATGAAATTTGAAAGATTGAGGCTGTGTGTGTGATGTGATTTGGATTTGTGCTGACACTG 60

QY 61 CCTGCAGAAATTAATTCAGGCACAAATCGTCACCGCAATTCATTTGGCAACACGATGGC 120  
 DB 61 ACGGCTGTGCGAGCTCATCGGAACGATTTATGATATATAGATAGGACACACAGCGGA 120

QY 121 TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTGGGACAAATGATTTCTCAATCATGGC 180  
 DB 121 TACGATTTTGAATTTTGAAGGATTTACGGAATACCTCG---ATGACACTCAATAACGGC 177

QY 181 GGTACGTTGAGTCCCAATGGAACAAATGTTTAAACAATATTTTCGTTAAAGTTAAAAA 240  
 DB 178 GGGGCAATTTAGTGCAGCTGGAACAAATTTGGAAATGCGCTTATTTGAAAGGAAAG 237

QY 241 TTCAATGAACACAAACACCAACAAAGTTGGTAAACATGTCATTAACACTACCGAGCAAC 300  
 DB 238 TTTGATTCACATAAAACTCATCACTTGGCAACATCTCCATCACTACACGAGCC 297

QY 301 TTCCAAACAAATGTAATCGGTATTTATGCGTCTATGTTGACTGTTGACCTCTTGTG 360  
 DB 298 TTTAACCCTGGCGGAATTCCTATTTATGTTCTATGCTGAGACAAATCTCCATTAGCT 357

QY 361 GAATATATATTTGTCAGAGTTGGGGCAACTGGCTCCACAGGAGCAACGCTTAAGGG 420  
 DB 358 GAATACTACATTTGTTGAGTTCATGGGGCACATATCGTCCAAACAGG---AACGTATAAGGA 414

QY 421 ACCATCACTGTTGATGGAGGACATATATATCTACGAGACTCTTAGAGTCAATCAACCC 480  
 DB 415 TCATTTTATGCGATGGAGGACATATGACATATATGAACGCTCGGTGCTCAATGAGCT 474

QY 481 TCATTAAGGGGATTCACCAATTTAAACAATATTTGAGTGTTCGAAGATCGAAACGACG 540  
 DB 475 TCTATCATTTGGAGACCTACCTTCAACAATATTTGGAGTGTAGTCAAAACAAACGACCA 534

QY 541 AGTGGACGATTTCTGTGACGACCACTTTAGAGCGTGGGAAACTTAGGATGATATG 600  
 DB 535 AGCGGAACCTGTTTCCGCTAGTGAGCATTTTAAATAATGGGAAAGCTTAGGCAATG 594

QY 601 GGGAAATGTATGAAGTTCGCTTACTGTAGAGGCTATCAAGATAGCGGAAGTGTAAAT 660  
 DB 595 GGAATAATGTATGAACACAGATTAATCTGTAGAGGCTACCGAAGCAACGGAAGTGGAT 654

QY 661 GTATATAGCAATACATAAGAAAT 684  
 DB 655 GTCATGACGAATCAGCTGATGATT 678

RESULT 11  
 ID AAQ80923 standard; DNA; 1022 BP.  
 AC AAQ80923;  
 AC AAQ80923;  
 DT 02-AUG-1995 (first entry)  
 DE B. pumilus xylanase gene.  
 KW Xylanase; endo-1,4-beta-D-xylanase; wood pulp; pulping; biobleaching;  
 KW bleaching; Bacillus licheniformis; ss.  
 OS Bacillus pumilus.  
 FH Key Location/Qualifiers  
 FT 5'UTR 1..185  
 FT /tag= a  
 FT /note= "claim 12"  
 FT misc\_difference 107  
 FT /tag= b  
 FT /note= "base n at position 107 is not identified in the

FT CDS specification"  
 FT 186..869  
 FT /\*tag= c  
 FT sig\_peptide 186..266  
 FT /\*tag= d  
 FT /note= "claim 11"  
 FT mat\_peptide 267..866  
 FT /\*tag= e  
 FT /EC number= "3.2.1.8"  
 FT /note= "claim 10"  
 XX GB2279955-A.  
 PN 18-JAN-1995.  
 PD 15-JUL-1993; 93GB-00014780.  
 PF 15-JUL-1993; 93GB-00014780.  
 PR (SOLV ) SOLVAY & CIE.  
 PA Amory A, Andre C, De Buyl E, Detroz R, Lahaye A, Ledoux P;  
 PI WPI; 1995-039214/06.  
 DR P-PSDB; AAR68849.  
 XX Purified xylanase from Bacillus pumilus PRL B12 - esp. produced in  
 PT transformed Bacillus licheniformis, and related DNA vectors, etc., used  
 PT for pre-treatment of wood pulp to reduce chlorine or ozone consumption in  
 PT subsequent bleaching.  
 XX Claim 13; Fig 1a-1b; 97pp; English.  
 XX A Bacillus pumilus PRL B12 (ATCC 55443) gene library was screened for  
 CC recombinant plasmids carrying the xylanase gene. A chromosomal fragment  
 CC obtained from isolate pBPX1 was subcloned and expressed in Escherichia  
 CC coli JM109. The sequence of the xylanase gene (as a 1022 bp Sau3AI  
 CC fragment) carried by a selected transformant is given in AAQ80923  
 XX Sequence 1022 BP; 329 A; 197 C; 229 G; 266 T; 0 U; 1 Other;  
 SQ Query Match 36.7%; Score 273; DB 2; Length 1022;  
 Best Local Similarity 63.4%; Pred. No. 1.8e-70;  
 Matches 452; Conservative 0; Mismatches 255; Indels 6; Gaps 2;  
 QY 1 ATGAGACAAAAGAAATTTGAGTGGTGTATTTAGCTTTTATGTTTGTGCTACCTTA 60  
 DB 186 ATGAAATTTGAAAGATTGAGGCTGTGTGTGATGTGATTTGGATTTGTGCTGACACTG 245

QY 61 CCTGCAGAAATTAATTCAGGCACAAATCGTCACCGCAATTCATTTGGCAACACGATGGC 120  
 DB 246 ACGGCTGTGCGAGCTCATCGGAACGATTTATGATATAGGATAGGACACACAGCGGA 305

QY 121 TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTGGGACAAATGATTTCTCAATCATGGC 180  
 DB 306 TACGATTTTGAATTTTGAAGGATTTACGGAATACCTCG---ATGACACTCAATAACGGC 362

QY 181 GGTACGTTGAGTCCCAATGGAACAAATGTTTAAACAATATTTTCGTTAAAGTTAAAAA 240  
 DB 363 GGGGCAATTTAGTGCAGCTGGAACAAATTTGAAATGCGCTTATTTGAAAGGAAAG 422

QY 241 TTCAATGAACACAAACACACAAAGTTGGTAAACATGTCCATAAATACACGAGCAAC 300  
 DB 423 TTTGATTCACATAAAACTCATCACTTGGCAACATCTCCATCACTACACGAGCC 482

QY 301 TTCCAAACAAATGTAATCGGTATTTATGCGTCTATGTTGACTGTTGACCTCTTGTG 360  
 DB 483 TTTAACCCTGGCGGAATTCCTATTTATGTTCTATGCTGAGACAAATCTCCATTAGCT 542

QY 361 GAATATATATTTGTCAGAGTTGGGGCAACTGGCTCCACAGGAGCAACGCTTAAGGG 420  
 DB 543 GATATCTACATTTGTTGAGTTCATGGGACATATCGTCCAAACAGG---AACGTATAAGGA 599



PT New recombinant DNA molecule comprising a sequence encoding feruloyl  
 PT esterase protein, useful for treating grasses and other plant materials  
 PT used in pulp and paper industries, feed processing and food additives.  
 XX  
 PS Disclosure; Page 97-99; 105pp; English.  
 XX  
 CC The present sequence is a DNA encoding xylanase A (XynA) from *Clostridium*  
 CC *stercorarium*. XynA contains family VI cellulose binding domain (CBD)  
 CC which is homologous to the CBD of xylanase Z (XynZ) of *Clostridium*  
 CC *thermocellum*. XynZ is an enzymatic component of *C. thermocellum*  
 CC cellulosome and has a multi-domain structure which includes a dockerin  
 CC domain, a catalytic xylanase domain, a family VI cellulose binding domain  
 CC and a domain of unknown function. The unknown domain in the N-terminal  
 CC region of XynZ has been found to contain feruloyl (phenolic acid)  
 CC esterase which is involved in the degradation of plant cell wall

RESULT 13	
AAT90972	
ID	AAT90972 standard; DNA; 1190 BP.
XX	
XX	
AC	AAT90972;
XX	.
DT	22-MAY-1998 (first entry)
XX	
DE	Nucleotide sequence encoding an enzyme with xylanase activity.
XX	
XX	xytB gene; xylanase; enzyme; Dictyoglomus thermophilum strain Rt46B.1;
KW	G-Xylanase; beta -1,4-xylanase activity; activity; thermal stability;
KW	biological bleaching; cellulose product; paper pulp; ss.
XX	
OS	Dictyoglomus thermophilum.
XX	.
XX	



Key	Location/Qualifiers
Key CDS	67..1149
FT	/*tag= a
FT	/product= "xylanase"
FT	67..138
FT	/*tag= b
FT	139..1146
FT	/*tag= c
XX	
PN	W09736995-A2.
XX	
PD	09-OCT-1997.
XX	
PF	27-MAR-1997; 97WO-NZ000042.
XX	
PR	29-MAR-1996; 96NZ-00286296.
XX	
PA	(PACI-) PACIFIC ENZYMES LTD.
XX	
PI	Bergquist PL, Gibbs MD, Morris D;
XX	
DR	WPI; 1997-503090/46.
DR	P-PSDB; AAW30267.
XX	
PT	Dictyoglomus thermophilum xylanase gene - used for producing thermally stable enzymes for the bleaching of cellulase products, especially paper pulp.
XX	
PS	Disclosure; Fig 3; 34pp; English.
XX	
CC	The present sequence represents a RT46B-1 xynB gene and encodes a xylanase enzyme. The present sequence is derived from a gene contained within Dictyoglomus thermophilum strain RT46B.1. The xylanase is contained within the family of enzymes known as G-Xylanases, and has beta -1,4-xylanase activity. The enzyme has high activity and high thermal stability with optimum activity at 85 degrees Celcius and pH 6.5. The xylanase enzyme is used for the biological bleaching of cellulose products, especially paper pulp. Use of the enzyme ensures that waste streams from the biological bleaching will include less toxic material
XX	
SQ	Sequence 1190 BP; 386 A; 197 C; 277 G; 330 T; 0 U; 0 Other;
Query Match 28.5%; Score 212.2; DB 2; Length 1190;	
Best Local Similarity 62.4%; Pred. No. 2.2e-52;	
Matches 369; Conservative 0; Mismatches 213; Indels 9; Gaps 2	
Qy	84 AATCGTCAACGCAATTCATTGGCAACCAACGATGGCTATGATTGATTTGGAAAGA 143
Db	150 AACACTAACAGTATGCAAGCGGTACTTTTGGTGGCTACTACTATGAACCTATGAAAGA 209
Qy	144 TAGCGGTGGCTCTGGGCAATGATTTCTCAATCATGGCGGTACGTTTCAGTGCCCAATGGAA 203
Db	210 TA----CAGGGAATACACATGACTGTATACACAGGAAGTTTGTCTGTCAGTGGAG 266
Qy	204 CAATGTTTAAACAATATTATTTCGTTAAAGGTAAGAAAATTTCAATGAACAACAACACCA 263
Db	267 CAATATAACAATGCATTATTTCAGAACAGGTAAAGATACAAC-----CAAAAACCTGGCA 320
Qy	264 ACAAGTTGGTAACATGTCATTAACATACGGAGCCAACTTCCAAACCAATGGTTAATGCCGA 323
Db	321 GTCATTAGGCACCTATTAGAATCACCTACTTCAGCCACATATAATCTCTAATGGTAACTCCTA 380
Qy	324 TTTATGCCCTATGGTTGGACTGTTTGACCCTCTGTGCAATATTATTTGTTCGACAGTTTG 383
Db	381 CTTATGTATCTATGGTTGGTCTCTAATTCCTTTAGTAGTTTTTACATTTGTAGAAAGTTG 440
Qy	384 GGGCAACTCGGGCTCCACAGGAGCAACGCTTAAGGGGACCATCCTCTTTGTATGGAGGAAC 443
Db	441 GGGTAATTTGGCGTCCACAGGTGCAACCTCTCTTGGACAGGTACTTATCGACGGTGGTAC 500
Qy	444 ATATGATATCTACGAGACTCTTTAGAGTCAATCAACCTCCATTAAGGGGATTTGCCACATT 503
Db	501 CTATGACATTTACAGCACTTACCCGTGTAAATCAGCCATCTATTGTGCGTACAGACTACTTT 560



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 17:09:58 ; Search time 85.4689 Seconds  
(without alignments)  
6187.359 Million cell updates/sec

Title: US-09-909-207-4  
Perfect score: 744  
Sequence: 1 ATGACACAAGAAATTGAC.....TAACTTTGGATATAAACCAAT 744

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/5A COMB.seq:  
2: /cgn2\_6/prodata/1/ina/5B COMB.seq:  
3: /cgn2\_6/prodata/1/ina/6A COMB.seq:  
4: /cgn2\_6/prodata/1/ina/6B COMB.seq:  
5: /cgn2\_6/prodata/1/ina/PCUTS COMB.seq:  
6: /cgn2\_6/prodata/1/ina/backfileseq1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	744	100.0	744	3	US-08-470-953A-4
2	744	100.0	744	3	US-08-470-953A-5
3	744	100.0	1513	3	US-08-470-953A-10
4	744	100.0	1513	3	US-08-470-953A-11
5	713.6	95.9	744	3	US-09-189-060B-1
6	713.6	95.9	871	1	US-08-698-978-1
7	671.6	90.3	744	3	US-08-501-126-18
8	663	89.1	663	3	US-08-470-953A-1
9	663	89.1	663	3	US-08-470-953A-2
10	587.2	78.9	744	3	US-09-189-060B-11
11	273	36.7	1022	3	US-08-275-526C-1
12	273	36.7	1022	3	US-08-275-526C-35
13	273	36.7	1022	4	US-09-076-677-1
14	273	36.7	1022	4	US-09-076-677-35
15	273	36.7	1022	4	US-09-073-055-1
16	273	36.7	1022	4	US-09-073-055-35
17	272.8	36.7	681	3	US-08-275-526C-30
18	272.8	36.7	681	3	US-08-275-526C-32
19	272.8	36.7	681	4	US-09-076-677-30
20	272.8	36.7	681	4	US-09-073-055-30
21	272.8	36.7	681	4	US-09-073-055-32
22	272.8	36.7	681	4	US-09-390-234-23
23	255.6	34.4	2364	3	US-09-603-311-23
24	255.6	34.4	2364	4	US-08-275-526C-26
25	255	34.3	600	3	US-08-275-526C-34
26	255	34.3	600	3	US-09-076-677-26
27	255	34.3	600	4	US-09-076-677-26

28	255	34.3	600	4	US-09-076-677-34
29	255	34.3	600	4	US-09-073-055-26
30	255	34.3	600	4	US-09-073-055-34
31	204.8	27.5	1244	3	US-08-591-685-12
32	154.4	20.8	164	3	US-08-501-126-15
33	154.4	20.8	164	3	US-08-501-126-16
34	132.4	17.8	294	3	US-08-817-946-4
35	132.4	17.8	294	4	US-09-639-354A-4
36	119.4	16.0	1375	2	US-08-468-812-1
37	119.4	16.0	1375	3	US-08-590-563-1
38	119.4	16.0	1375	4	US-09-770-621-1
39	119.4	16.0	1375	4	US-09-233-832-1
40	119	16.0	229	3	US-09-189-060B-53
41	118.4	15.9	573	1	US-08-709-912-18
42	118.4	15.9	573	2	US-09-047-370-18
43	117.8	15.8	1207	1	US-08-575-964-2
44	117.8	15.8	1207	1	US-08-963-500-2
45	115.2	15.5	213	3	US-09-189-060B-49

ALIGNMENTS

RESULT 1  
US-08-470-953A-4  
; Sequence 4, Application US/08470953A  
; Patent No. 6346407

; GENERAL INFORMATION:

; APPLICANT: ANDREE LAHAYE

; APPLICANT: ERIC DE BUYL

; APPLICANT: PIERRE LEDOUX

; APPLICANT: RENE DETROZ

; TITLE OF INVENTION: xylanase, microorganisms produced it.

; TITLE OF INVENTION: DNA molecule, processes for preparation of this xylanase

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE

; STREET: 2000 K St., N.W., Suite 200

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20006

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/470,953A

; FILING DATE: 6-OCTOBER-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Wilhem F. Gadiano, Esq.

; REGISTRATION NUMBER: 37,136

; REFERENCE/DOCKET NUMBER: 4121-40

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-429-0625

; TELEFAX: (202) 293-1850

; TELEX: 650 383-5605

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 744 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; ORIGINAL SOURCE:

; STRAIN: Bacillus

US-08-470-953A-4





TITLE OF INVENTION: Xylanase, microorganisms produced it,  
TITLE OF INVENTION: DNA molecule, processes for preparation of this xylanase  
NUMBER OF INVENTION: and uses thereof  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
STREET: 2000 K St., N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/470,953A  
FILING DATE: 6-OCTOBER-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilhem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-40  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1513 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: Bacillus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 620...1363  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 701...1363  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 620...700

Db 860 TTCAATGAACACAAACACACACCAAGTTGGTAACATGTCCTCATAAAATACGGAGCAAC 919  
Qy 301 TTCCAAACCAAAAGTAAATCGTATTTATCGGTCTATGGTTGACTGCTTGCACCTCTTGTC 360  
Db 920 TTCCAAACCAAAAGTAAATCGTATTTATCGGTCTATGGTTGACTGCTTGCACCTCTTGTC 979  
Qy 361 GAATATATATATGTCGACAGTGTGGGCAACTGGCGTCCACAGGAGCAACGCCCTAAGGGG 420  
Db 980 GAATATATATGTCGACAGTGTGGGCAACTGGCGTCCACAGGAGCAACGCCCTAAGGGG 1039  
Qy 421 ACCATCACTGTTGATGGAGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480  
Db 1040 ACCATCACTGTTGATGGAGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 1099  
Qy 481 TCATTAAGGGGATTCGCACATTTAAACAATATTGGAGTGTTCGAAGATCGAAGCCACG 540  
Db 1100 TCATTAAGGGGATTCGCACATTTAAACAATATTGGAGTGTTCGAAGATCGAAGCCACG 1159  
Qy 541 AGTGGCACGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 600  
Db 1160 AGTGGCACGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 1219  
Qy 601 GGGAAAATGTATGAAGTCGCGCTTACTGTAGAAAGGCTATCAAAAGTAGCGGAAGTCTAAT 660  
Db 1220 GGGAAAATGTATGAAGTCGCGCTTACTGTAGAAAGGCTATCAAAAGTAGCGGAAGTCTAAT 1279  
Qy 661 GTATATAGCAATACACTTAAGATTAACGGTAACCCCTCTCTCAACTATTAGTAAATGACGAG 720  
Db 1280 GTATATAGCAATACACTTAAGATTAACGGTAACCCCTCTCTCAACTATTAGTAAATGACGAG 1339  
Qy 721 AGCATAACTTTGGATAAAACAAT 744  
Db 1340 AGCATAACTTTGGATAAAACAAT 1363

RESULT 5  
US-09-189-060B-1  
; Sequence 1, Application US/09189060B  
; Patent No. 6270968  
; GENERAL INFORMATION:  
; APPLICANT: Dalboge, Henrik  
; APPLICANT: Sandal, Thomas  
; APPLICANT: Kauppinen, Markus  
; APPLICANT: Borge, Diderichsen  
; TITLE OF INVENTION: Method Of Providing No. 6270968el DNA Sequences  
; FILE REFERENCE: 4772.204-US  
; CURRENT APPLICATION NUMBER: US/09/189,060B  
; CURRENT FILING DATE: 1998-11-10  
; PRIOR APPLICATION NUMBER: PCT/DK97/00216  
; PRIOR FILING DATE: 1997-05-12  
; NUMBER OF SEQ ID NOS: 74  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 744  
; TYPE: DNA  
; ORGANISM: Bacillus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(744)  
US-09-189-060B-1

Query Match 95.9%; Score 713.6; DB 3; Length 744;  
Best Local Similarity 97.4%; Pred. No. 2.9e-213;  
Matches 725; Conservative 0; Mismatches 19; Indels 0; Gaps 0;  
Qy 1 ATGAGACAAAAGAAATTCAGCGTTCGATTTTACGCTTTTGTAGTTTGTGCTAACTTA 60  
Db 1 ATGAGACAAAAGAAATTCAGCGTTCGATTTTACGCTTTTGTAGTTTGTGCTAACTTA 60  
Qy 61 CCTGCAGAAAATAATTCAGGCAAAATCGTACCGACAAATTCATTCGCAACACGATGGC 120  
Db 61 CCTGCAGAAAATAATTCAGGCAAAATCGTACCGACAAATTCATTCGCAACACGATGGC 120

Query Match 100.0%; Score 744; DB 3; Length 1513;  
Best Local Similarity 100.0%; Pred. No. 1.3e-222;  
Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGAGACAAAAGAAATTCAGCGTTCGATTTTACGCTTTTGTAGTTTGTGCTAACTTA 60  
Db 620 ATGAGACAAAAGAAATTCAGCGTTCGATTTTACGCTTTTGTAGTTTGTGCTAACTTA 679  
Qy 61 CTGCGAGAAATAATTCAGGCAAAATCGTACCGACAAATTCATTCGCAACACGATGGC 120  
Db 680 CTGCGAGAAATAATTCAGGCAAAATCGTACCGACAAATTCATTCGCAACACGATGGC 739  
Qy 121 TATGATTATGAATTTTGGAAAGATAGCGGTGCTCTGGACAAATGATCTCAATCATGCG 180  
Db 740 TATGATTATGAATTTTGGAAAGATAGCGGTGCTCTGGACAAATGATCTCAATCATGCG 799  
Qy 181 GGTACGTTTCAGTGCCCAATGGAACAAATGTTAAACAATATTTCGTAAGGTAATAAAA 240  
Db 800 GGTACGTTTCAGTGCCCAATGGAACAAATGTTAAACAATATTTCGTAAGGTAATAAAA 859  
Qy 241 TTCAATGAACACAAACACACCAAGTTGGTAACATGTCCTAATAACTACCGAGCAAC 300

```

121  TATGATTATGAATTTGGAAAGATAGCGTGGCTCTGGGACAAATGATTTCTCAATCATGCG 180
121  TATGATTATGAATTTGGAAAGATAGCGTGGCTCTGGGACAAATGATTTCTCAATCATGCG 180
181  GGTAGCTTCAGTGGCCCAATGGAACAATGTTAAACAACATATTTCCGTAAGGTAAGGTA 240
181  GGTAGCTTCAGTGGCCCAATGGAACAATGTTAAACAACATATTTCCGTAAGGTAAGGTA 240
241  TTCAATGAAACACAAACACACCAACAAAGTTGGTAACATGTCATATACTACGAGCCAAC 300
241  TTCAATGAAACACAAACACACCAACAAAGTTGGTAACATGTCATATACTACGAGCCAAC 300
301  TTCCACCAAAATGGAATGCGTATTTATCGCTTATGTTGACATGTTGACCTCTTGTG 360
301  TTCCACCAAAATGGAATGCGTATTTATCGCTTATGTTGACATGTTGACCTCTTGTG 360
361  GAATATTATTTGTCGACATGTTGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGG 420
361  GAATATTATTTGTCGATAGTTGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGG 420
421  ACATCAGCTGTTGATGGAGGAACATATGATATCTAGGACATCTTAGAGTCAATCAACCC 480
421  ACCATCAGCTGTTGATGGAGGAACATATGATATCTAGGACATCTTAGAGTCAATCAACCC 480
481  TCCATTAAAGGGATGCGCAATTTAAACAATTTGAGGTTGCGAAGATCGAAACGCAAG 540
481  TCCATTAAAGGGATGCGCAATTTAAACAATTTGAGGTTGCGAAGATCGAAACGCAAG 540
541  AGTGCGACCATTTCTGTCAGCAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 600
541  AGTGCGACCATTTCTGTCAGCAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 600
601  GGGAAATGATGAAGTCGCGTCTACTGTAGAGGCTTCAAGAGTACGAGGAGTGTAT 660
601  GGGAAATGATGAAGTCGCGTCTACTGTAGAGGCTTCAAGAGTACGAGGAGTGTAT 660
661  GTATATAGCAATACATGAATTAACGGTAAACGGTAAACGGTAAACGGTAAACGGTAA 720
661  GTATATAGCAATACATGAATTAACGGTAAACGGTAAACGGTAAACGGTAAACGGTAA 720
721  AGCAATACTTTGGTAAACCAAT 744
721  AGCAATACTTAGATAAACAAT 744

```

```

RESULT 6
US-08-698-978-1
; Sequence 1, Application US/08698978
; Patent No. 5770424
; GENERAL INFORMATION:
; APPLICANT: Outtrup, Helle
; APPLICANT: Danbmann, Claus
; APPLICANT: Olsen, Arne Agerlin
; APPLICANT: Bisg rd-Frantzen, Henrik
; APPLICANT: Sch lein, Martin
; APPLICANT: Jorgensen, Per Linnaa
; TITLE OF INVENTION: DNA CONSTRUCTS AND METHODS OF PRODUCING
; TITLE OF INVENTION: XYLANOLYTIC ENZYMES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5770424 of No. 5770424 disk of No. 5770424th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/698,978

```

```

; FILING DATE: 16 August 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/470,398
; FILING DATE: 06-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/343,600
; FILING DATE: 30-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/DK93/00218
; FILING DATE: 2-JUL-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Gregg, Valeta A.
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 3794,400-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 871 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORGANISM: Bacillus agaradherens
; STRAIN: NCIMB 40482
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 82..744
; US-08-698-978-1
;
; Query Match 95.9%; Score 713.6; DB 1; Length 871;
; Best Local Similarity 97.4%; Pred. No. 3.2e-213;
; Matches 725; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
;
QY 1 ATGAGACAAAAGAAATGACGTGTTGATTTTAGCCTTTTATAGTCTTTTATAGTCTTAACTTA 60
DB 1 ATGAGACAAAAGAAATGACGTGTTGATTTTAGCCTTTTATAGTCTTTTATAGTCTTAACTTA 60
QY 61 CTGAGCAAAATTAATTCAGGCAAAATTCGTACCGCAAAATTCATTCGTAACCAACGATGCG 120
DB 61 CTGAGCAAAATTAATTCAGGCAAAATTCGTACCGCAAAATTCATTCGTAACCAACGATGCG 120
QY 121 TATGATTATGAATTTTGGAAAGATAGCGTGGCTCTGGGACAAATGATTTCTCAATCATGCG 180
DB 121 TATGATTATGAATTTTGGAAAGATAGCGTGGCTCTGGGACAAATGATTTCTCAATCATGCG 180
QY 181 GGTAGCTTCAGTGGCCCAATGGAACAATGTTAAACAACATATTTCCGTAAGGTAAGGTA 240
DB 181 GGTAGCTTCAGTGGCCCAATGGAACAATGTTAAACAACATATTTCCGTAAGGTAAGGTA 240
QY 241 TTCAATGAAACACAAACACACCAACAAAGTTGGTAACATGTCATATACTACGAGCCAAC 300
DB 241 TTCAATGAAACACAAACACACCAACAAAGTTGGTAACATGTCATATACTACGAGCCAAC 300
QY 301 TTCCACCAAAATGGAATGCGTATTTATCGCTTATGTTGACATGTTGACCTCTTGTG 360
DB 301 TTCCACCAAAATGGAATGCGTATTTATCGCTTATGTTGACATGTTGACCTCTTGTG 360
QY 361 GAATATTATTTGTCGACATGTTGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGG 420
DB 361 GAATATTATTTGTCGATAGTTGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGG 420
QY 421 ACCATCAGCTGTTGATGGAGGAACATATGATATCTAGGACATCTTAGAGTCAATCAACCC 480
DB 421 ACCATCAGCTGTTGATGGAGGAACATATGATATCTAGGACATCTTAGAGTCAATCAACCC 480
QY 481 TCCATTAAAGGGATGCGCAATTTAAACAATTTGAGGTTGCGAAGATCGAAACGCAAG 540
DB 481 TCCATTAAAGGGATGCGCAATTTAAACAATTTGAGGTTGCGAAGATCGAAACGCAAG 540

```



541 AGTGGCAGATTCTGTGACCAACACCTTTAGAGCGTGGGAAACCTTAGGATGAATATG 600  
 541 AGTGGCAGAAATTTCTGTGACCAACACCTTTAGAGCGTGGGAAACCTTAGGATGAATATG 600  
 601 GGGAAATGATGATGAAGTGGCGCTTACTGTAGAGGCTATCAAGTAGCGGAAGTGCTAAT 660  
 601 GGGAAATGATGATGAAGTGGCGCTTACTGTAGAGGCTATCAAGTAGCGGAAGTGCTAAT 660  
 661 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTAATGACAG 720  
 661 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTAATGACAG 720  
 721 AGCATAACTTTGGTAAATAAACAAAT 744  
 721 AGCATAACTTTAGATAAAAAACAAT 744

RESULT 7  
 US-08-501-126-18  
 ; Sequence 18, Application US/08501126  
 ; Patent No. 6140095  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Van Solingen, Pieter  
 ; APPLICANT: Williams, Diane P.  
 ; APPLICANT: Iverson, Sara  
 ; APPLICANT: Farrell, Roberta L.  
 ; APPLICANT: Herbes, Wilhelmina T.  
 ; APPLICANT: Van Der Kleij, Wilhelmus A.  
 ; APPLICANT: Herweijer, Margaretha A.  
 ; APPLICANT: Van Beckhoven W.C., Rudolf F.  
 ; APPLICANT: Quax, Wilhelmus J.  
 ; APPLICANT: Jones, Brian E.  
 ; TITLE OF INVENTION: ALKALI-TOLERANT XYLANASES  
 ; NUMBER OF SEQUENCES: 20  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESS: Morrison & Foerster  
 ; STREET: 2000 Pennsylvania Avenue, NW  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20006-1812  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/501.126  
 ; FILING DATE: 29-DEC-1995  
 ; CLASSIFICATION: 425  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Muraishige, Kate H.  
 ; REGISTRATION NUMBER: 29,959  
 ; REFERENCE/DOCKET NUMBER: 4615-0057.00  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 887-1500  
 ; TELEFAX: (202) 887-0763  
 ; INFORMATION FOR SEQ ID NO: 18:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 744 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; STRAIN: 1-43-3  
 ; INDIVIDUAL ISOLATE: CBS672.93  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..744  
 ; OTHER INFORMATION: /product= "xylanase"

US-08-501-126-18  
 Query Match 90.3%; Score 671.6; DB 3; Length 744;  
 Best Local Similarity 95.7%; Pred. No. 4e-200;  
 Matches 712; Conservative 0; Mismatches 29; Indels 3; Gaps 2;  
 QY 1 ATGACAGAAAAAGAAATGACGTTGATTTTAGCCCTTTTATAGTTTGTGTTTGTGACCTTA 60  
 DB 1 ATGACGAAAAAGAAATGACGTTGA--TTAACCTTTTATAGTTTGTGACCTTAACCTTA 57  
 QY 61 CTGCGAGAAATATTCAGGCACAAATCGTCACCGACAATTCCTATGGCAACGATGGC 120  
 DB 58 CTGCGAAGATAAGTCAGGCACAAATCGTCACCGACAATTCCTATGGCAACGATGGC 117  
 QY 121 TATGATTATGAATTTTGGAAAGATAGCGTGGCTCTGGGACAATGATTCCTCAATCATGGC 180  
 DB 118 TATGATTATGAATTTTGGAAAGATAGCGTGGCTCTGGGACAATGATTCCTCAATCATGGC 177  
 QY 181 GGTACGTTTCAGTCCCAATGGAACAATGTTTAAACAATATTTTCCGTAAAGGTAAAAA 240  
 DB 178 GGTACGTTTCAGTCCCAATGGAATATGTTTAAACAATATTTTCCGTAAAGGTAAAAA 237  
 QY 241 TTCAATGAAAAACACAAACACACACCAACCAAGTTGGTAACATGTCCTAATAACTACGAGCAAC 300  
 DB 238 TTCAATGAAAAACACAAACACACCAACCAAGTTGGTAACATGTCCTAATAACTACGAGCAAC 297  
 QY 301 TTCCAACCAATGTAATGCTATTTATGCTCTATGCTTGGACTGTGTGACCCCTCTTGTG 360  
 DB 298 TTCCAGCCAAACGGTAATGCTATTTATGCTCTATGCTTGGACTGTGTGACCCCTCTTGTG 357  
 QY 361 GAATATTATATTTGTCGACAGTTGGGGCAACTGGGCTCCACAGAGCAACGCTTAAGGG 420  
 DB 358 GAATATTATATTTGTCGACAGTTGGGGCAACTGGGCTCCACAGAGCAACGCTTAAGGG 417  
 QY 421 ACCATCACTGTTGATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480  
 DB 418 ACCATCACTGTTGATGAGGAAACATATGATATCTATGAAACTCTTAGAGTCAATCAACCC 477  
 QY 481 TCCATTAAGGGGATTGGCACAATTTAAACAATTTTGAAGTGTTCGAAGATCGAAACGACG 540  
 DB 478 TCCATTAAGGGGATTGGCACAATTTAAACAATTTTGAAGTGTTCGAAGATCGAAACGACG 537  
 QY 541 AGTGGCAGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAACCTTAGGATGAATATG 600  
 DB 538 AGTGGCACAATTTCTGTGACCAACCACTTTAGAGCGTGGGAAACCTTAGGATGAATATG 597  
 QY 601 GGGAAATGATGAAGTGGCGCTTACTGTAGAGGCTATCAAGTAGCGGAAGTGCTAAT 660  
 DB 598 GGGAAATGATGAAGTGGCGCTTACTGTAGAGGCTATCAAGTAGCGGAAGTGCTAAT 657  
 QY 661 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTAATGACAG 720  
 DB 658 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTAATGACAG 717  
 QY 721 AGCATAACTTTGGTAAATAAACAAAT 744  
 DB 718 AGCATAACTTTAGATAAAAAACAAT 741

RESULT 8  
 US-08-470-953A-1  
 ; Sequence 1, Application US/08470953A  
 ; Patent No. 6346407  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ANDREE LAHAVE  
 ; APPLICANT: ERIC DE BUYL  
 ; APPLICANT: PIERRE LEDOUX  
 ; APPLICANT: FIERNE DETROZ  
 ; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
 ; TITLE OF INVENTION: DNA molecule, processes for preparation of this xylanase  
 ; TITLE OF INVENTION: and uses thereof  
 ; NUMBER OF SEQUENCES: 29  
 ; CORRESPONDENCE ADDRESS:



ADDRESSER: WILLIAN BRINKS HOFER GILSON & LIONE  
 STREET: 2000 K St., N.W., Suite 200  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20006  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/470,953A  
 FILING DATE: 6-OCTOBER-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Wilhem F. Gadiano, Esq.  
 REGISTRATION NUMBER: 37,136  
 REFERENCE/DOCKET NUMBER: 4121-40  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-429-0625  
 TELEFAX: (202) 293-1850  
 TELEX: 650 383-5605  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 663 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 ORIGINAL SOURCE:  
 STRAIN: Bacillus  
 US-08-470-953A-1

Query Match 89.1%; Score 663; DB 3; Length 663;

Best Local Similarity 100.0%; Pred. No. 1.9e-197; Indels 0; Gaps 0;  
Matches 663; Conservative 0; Mismatches 0;

QY	82	CAAAATCGTCACCGACAAATTCATTTGGCAACACGAGTGGTATGATTAATGAAATTTTGAAA	141
DB	1	CAAAATCGTCACCGACAAATTCATTTGGCAACACGAGTGGTATGATTAATGAAATTTTGAAA	60
QY	142	GATAGCGGTGGCTCTGGGACAATGATTCATATGCGGGTACGTTTCAGTGGCCCAATGG	201
DB	61	GATAGCGGTGGCTCTGGGACAATGATTCATATGCGGGTACGTTTCAGTGGCCCAATGG	120
QY	202	AACAATGTTAAACAACATATATTCCTGAAGGTAAATAATTCATGAACACACAAACACAC	261
DB	121	AACAATGTTAAACAACATATATTCCTGAAGGTAAATAATTCATGAACACACAAACACAC	180
QY	262	CAACAAGTTGGTAACATGTCATAAACTACGAGCAAACTTCCAAACAAATGTAATGGG	321
DB	181	CAACAAGTTGGTAACATGTCATAAACTACGAGCAAACTTCCAAACAAATGTAATGGG	240
QY	322	TATTTATGCGTCTATGGTTGAGTGTGACCCCTCTTGTGCAATATATATTTGCGACAGT	381
DB	241	TATTTATGCGTCTATGGTTGAGTGTGACCCCTCTTGTGCAATATATATTTGCGACAGT	300
QY	382	TGGGGCACTGGGCTCCACGAGCAACGCTTAAGGGACCATCACTGTTGATGGGGA	441
DB	301	TGGGGCACTGGGCTCCACGAGCAACGCTTAAGGGACCATCACTGTTGATGGGGA	360
QY	442	ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTCGCACA	501
DB	361	ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTCGCACA	420
QY	502	TTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACGAGTGGCGACGATTTCTGTGACG	561
DB	421	TTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACGAGTGGCGACGATTTCTGTGACG	480

QY	562	AACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATGGGAAATATGATGAATCGCG	621
DB	481	AACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATGGGAAATATGATGAATCGCG	540
QY	622	CTTACTGTAGAAGGCTATCAAGTAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA	681
DB	541	CTTACTGTAGAAGGCTATCAAGTAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA	600
QY	682	ATTAACGGTAACCCCTCTCTCAACTATTAGTAATGACGAGAGCATAACTTTGGATAAAAAAC	741
DB	601	ATTAACGGTAACCCCTCTCTCAACTATTAGTAATGACGAGAGCATAACTTTGGATAAAAAAC	660
QY	742	AAT 744	
DB	661	AAT 663	

# RESULT 9

US-08-470-953A-2  
 ; Sequence 2, Application US/08470953A  
 ; Patent No. 6346407  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ANDREE LAHAYE  
 ; APPLICANT: ERIC DE BUYL  
 ; APPLICANT: PIERRE LEDOUX  
 ; APPLICANT: RENE DETROZ  
 ; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
 ; TITLE OF INVENTION: DNA molecule, processes for preparation of this xylanase  
 ; NUMBER OF SEQUENCES: 29  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSER: WILLIAN BRINKS HOFER GILSON & LIONE  
 ; STREET: 2000 K St., N.W., Suite 200  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20006  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/470,953A  
 ; FILING DATE: 6-OCTOBER-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Wilhem F. Gadiano, Esq.  
 ; REGISTRATION NUMBER: 37,136  
 ; REFERENCE/DOCKET NUMBER: 4121-40  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-429-0625  
 ; TELEFAX: (202) 293-1850  
 ; TELEX: 650 383-5605  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 663 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; ORIGINAL SOURCE:  
 ; STRAIN: Bacillus  
 ; FEATURE:  
 ; NAME/KEY: mat\_peptide  
 ; LOCATION: 1..663  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..663  
 ; US-08-470-953A-2

```

Query Match      89.1%; Score 663; DB 3; Length 663;
Best Local Similarity 100.0%; Pred. No. 1.9e-197;
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 CAAATCGTCACCGACAAATCCATTTGGCAACACGATGGCTATGATTAATGAAATTTTGGAAA 141
DB 1 CAAATCGTCACCGACAAATCCATTTGGCAACACGATGGCTATGATTAATGAAATTTTGGAAA 60
QY 142 GATAGCGGTGCTCTGGGCAATGATTTCTCAATCATGCGGTGACGTTGAGTCCCAATGG 201
DB 61 GATAGCGGTGCTCTGGGCAATGATTTCTCAATCATGCGGTGACGTTGAGTCCCAATGG 120
QY 202 AACATGTTAAACAATATTTTCGTAAGAGTAAATAATTCATTAACAACAACAACACAC 261
DB 121 AACATGTTAAACAATATTTTCGTAAGAGTAAATAATTCATTAACAACAACAACACAC 180
QY 262 CAACAAGTTGTAACATGTCCTATAAATACGAGGCAACTTCCAAACCAAAATGGTAATGCG 321
DB 181 CAACAAGTTGTAACATGTCCTATAAATACGAGGCAACTTCCAAACCAAAATGGTAATGCG 240
QY 322 TATTTATGCGTCTATGTTGGACTGTTGACCCCTCTTGTGCAATATATATTTGTCACAGT 381
DB 241 TATTTATGCGTCTATGTTGGACTGTTGACCCCTCTTGTGCAATATATATTTGTCACAGT 300
QY 382 TGGGGCAACTGGCGTCCACGAGGCAAGCCCTAAGGGGACCATCACTGTTGATGGAGGA 441
DB 301 TGGGGCAACTGGCGTCCACGAGGCAAGCCCTAAGGGGACCATCACTGTTGATGGAGGA 360
QY 442 ACATATGATATACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATGGCCACA 501
DB 361 ACATATGATATACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATGGCCACA 420
QY 502 TTTAAACAATATGAGAGTTCGAGATCGAAGACGAAACGACGATGTCGTCAGC 561
DB 421 TTTAAACAATATGAGAGTTCGAGATCGAAGACGAAACGACGATGTCGTCAGC 480
QY 562 AACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATGGGAAATATGTAAGTCCG 621
DB 481 AACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATGGGAAATATGTAAGTCCG 540
QY 622 CTTACTAGTAAGGCTATCAAAAGTAGGGAAGTGTCTAATATATATAGCAATACACTAAGA 681
DB 541 CTTACTAGTAAGGCTATCAAAAGTAGGGAAGTGTCTAATATATATAGCAATACACTAAGA 600
QY 682 ATTAACGGTAACCTCTCTCAACTATTAGTATGACGAGACATACCTTTCGATAAAGAAC 741
DB 601 ATTAACGGTAACCTCTCTCAACTATTAGTATGACGAGACATACCTTTCGATAAAGAAC 660
QY 742 AAT 744
DB 661 AAT 663

```

```

RESULT 10
US-09-189-060B-11
; Sequence 11, Application US/09189060B
; Patent No. 6270968
; GENERAL INFORMATION:
; APPLICANT: Dalboge, Henrik
; APPLICANT: Sandal, Thomas
; APPLICANT: Kauppinen, Markus
; APPLICANT: Borge, Diderichsen
; TITLE OF INVENTION: Method Of Providing No. 6270968el DNA Sequences
; FILE REFERENCE: 4772.204-US
; CURRENT APPLICATION NUMBER: US/09/189,060B
; CURRENT FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/DK97/00216
; PRIOR FILING DATE: 1997-05-12
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 744

```

```

; TYPE: DNA
; ORGANISM: Hybrid
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(744)
; US-09-189-060B-11

```

```

Query Match      78.9%; Score 587.2; DB 3; Length 744;
Best Local Similarity 86.8%; Pred. No. 1e-173;
Matches 646; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 1 ATGAGACAAAAGAAATTTGACGTTGATTTTACCCCTTTTGTGTTTGTGCACTAACCTTA 60
DB 1 ATGAGACAAAAGAAATTTGACGTTGATTTTACCCCTTTTGTGTTTGTGCACTAACCTTA 60
QY 61 CCTCAGAGAAATTAATTCAGGCAAAATCGTCACCGACAAATTCATTTGGCAACACGATGGC 120
DB 61 CCTCAGAGAAATTAATTCAGGCAAAATCGTCACCGACAAATTCATTTGGCAACACGATGGC 120
QY 121 TATGATTATGAAATTTTGGAAAGATAGCGGTGGCTCTGGGACAAATGATTTCTCAATCATGGC 180
DB 121 TATGATTATGAAATTTTGGAAAGATAGCGGTGGCTCTGGGACAAATGATTTCTCAATCATGGC 180
QY 181 GGTACGTTCACTGCGCCCAATGGAAACAATGTTAAACAATATTTATTCGTAAGGTAAAGAAA 240
DB 181 GGTACGTTCACTGCGCCCAATGGAAACAATGTTAAACAATATTTATTCGTAAGGTAAAGAAA 240
QY 241 TTCAATGAAACACAAAACACACCAACCAAGTTGGTAAACATGTCATATAAATCTAGCGAGCAAC 300
DB 241 TTCAATGAAACACAAAACACACCAACCAAGTTGGTAAACATGTCATATAAATCTAGCGAGCAAC 300
QY 301 TTCCAAACCAATGTAATGCTATTTATGCTGCTATGCTGTTGGACTGTTGACCTCTTTGTC 360
DB 301 TTCCAAACCAATGTAATGCTATTTATGCTGCTATGCTGTTGGACTGTTGACCTCTTTGTC 360
QY 361 GAATATTATATTTGTCGAGTGTGGGCAACTGGGCTCCACGAGGACCAAGCCCTTAAGGGG 420
DB 361 GAATATTATATTTGTCGAGTGTGGGCAACTGGGCTCCACGAGGACCAAGCCCTTAAGGGG 420
QY 421 ACCATCACTGTTGATGAGGAAACATATGATATCTACAGACTCTTAGAGTCAATCAACCC 480
DB 421 ACCATCACTGTTGATGAGGAAACATATGATATCTACAGACTCTTAGAGTCAATCAACCC 480
QY 481 TCCATTAAAGGGATGTCACATTTAAACAATATTTGAGTGTTCGAAGATCGAAACGCAACG 540
DB 481 TCTATTGAGGACCGCCACCTTCAATCAGTACTGTCGATTCGACAGCAACGCGAC 540
QY 541 AGTGGCACGATTTCTGTGACCAACCACTTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 600
DB 541 AGCGGCACTGTCACACTACCGCAAAACCACTTTTAATGCTGGGCTGCTCTTTGGCATGATG 600
QY 601 GGGAAATGATGAAATGCGGCTTACTGTAGAAGGCTATCAAGTAGCGGAGTCTAAT 660
DB 601 GGTGCAATTCATTAATACCAAGATCTCTGTTACTGAGGGCTACCAATCTACCGAAGTCTAAT 660
QY 661 GTATATAGCAATACACTTAAGAATTAACGGTAAACCTCTCTCAACTATTAGTAAATGACGAG 720
DB 661 GTATATAGCAATACACTTAAGAATTAACGGTAAACCTCTCTCAACTATTAGTAAATGACGAG 720
QY 721 AGCATAACTTTGGATAAAACAAT 744
DB 721 AGCATAACTTTGGATAAAACAAT 744

```

```

RESULT 11
US-08-275-526C-1
; Sequence 1, Application US/08275526C
; Patent No. 6180382
; GENERAL INFORMATION:
; APPLICANT: DE BUYL, ERIC
; APPLICANT: LAHAYE, ANDR E
; APPLICANT: LEDOUX, PIERRE
; APPLICANT: AMORY, ANTOINE

```



Query Match 36.7%; Score 273; DB 3; Length 1022;  
 Best Local Similarity 63.4%; Pred. No. 2.4e-75;  
 Matches 452; Conservative 0; Mismatches 255; Indels 6; Gaps 2;

QY 1 ATGAGACAAAAGAAATTCAGCTTGATTTAGCTTTTATGTTTGTGCTACTAACCTTA 60  
 DB 186 ATGAATTTGAAGAAGATTGAGCTGTGTTGTGATGTATGATTTGTCTGACACTG 245

QY 61 CCTCAGAAATAATTCAGGCAAAATGCTACCGACAATTCATTCGCAACACAGATGGC 120  
 DB 246 ACGGCTGTGCGGCTCATGCGAAACGATTTATGATATAGGATAGGACACACAGCGGA 305

QY 121 TATGATATGATTTTGAAGAGATAGCGTGGCTCTGGACAATGATTTCTCAATCAGGC 180  
 DB 306 TACGATTTGAAATTTATGGAAGATTACGGAATACCTCG---ATGACACTCAATAACGCG 362

QY 181 GGTACGTTCACTGCGCAATGGAACAATGTTTAAACAATATTTATTCGTAAGGTAAGAAA 240  
 DB 363 GGGCAATTTAGTCAAGCTGGAACAATATTGGAATGCTTTATTCGAAAAGGAAGAAG 422

QY 241 TTCAATGAACACAAACACACACCAATGTTGGTAACATGTCCTAATAACTACGAGCCAAC 300  
 DB 423 TTTGATTCACATAAACTCATCACTTGGCAACATCTCCATCAACTCAACGACGCC 482

QY 301 TTCCAAACCAATGTAATGCGTATTTATGCGTCTATGTTGGAATGTTGACCTCTTGTG 360  
 DB 483 TTTAACCGGCGGGAATTCCTATTTATGTTGCTATGCTGGACACATCTCCATAGCT 542

QY 361 GAATATTATATTTGTCAGAGTTGGGCAACTGGCGTCCACAGGAGCAACGCCCTAAGGGG 420  
 DB 543 GAATATTATATTTGAGTTCATGGGCACTATCTGTCACAGG---AACGTATTAAGGA 599

QY 421 ACCATCACTGTTGATGGAGGACATATGATATCTACGAGACTCTTAGAGCTCAATCAACC 480  
 DB 600 TCATTTTATGCGGATGGAGGACATATGACATATATGAACGCTCCGCTGTCATCAGCCT 659

QY 481 TCCATTAGGGAATGGCCACATTTAAACAATATTGGAGTGTTCGAAGATCGAAGCAGC 540  
 DB 660 TCTATCATTTGAGAGCGTACCTTTCAACAATATTGGAGTGTACGTCGAACAACAGCACA 719

QY 541 AGTGCAGGATTTCTGTGAGCAACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATG 600  
 DB 720 AGCGAAACGGTCTCGTCACTGAGCATTTTAAATAATGGAAAGCTTAGGCGATGCAATG 779

QY 601 GGGAAATGATGAAGTCCGCTTACTGTAGAGGCTATCAAGTAGCGGAAGTCTTAAT 660  
 DB 780 GGAATAATGATGAACAGCATTAACCTGTAGAAGGCTACCGAAGCAACGGAAGTGCAGAT 839

QY 661 GTATATAGCAATACACTAAGATTAACGTTAACCGTACCTCTCACTATTAGTAA 713  
 DB 840 GTCATGAGCAATCAGCTGATGATTCGATAAAGCATATGAATAAGCCAGCA 892

RESULT 13  
 US-09-076-677-1  
 ; Sequence 1, Application US/09076677  
 ; Patent No. 6423523  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DE BUYL, ERIC  
 ; LAHAYE, ANDREE  
 ; LEDOUX, PIERRE  
 ; AMORY, ANTOINE  
 ; DETROZ, RENE  
 ; ANDRE, CHRISTOPHE  
 ; VETTER, ROMAN  
 ;  
 ; TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,  
 ; EXPRESSION VECTORS FOR SUCH XYLANASE AND  
 ; OTHER PROTEINS, HOST ORGANISMS THEREFOR AND  
 ; USE THEREOF  
 ;  
 ; NUMBER OF SEQUENCES: 35  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: WILLIAM BRINKS HOPER GILSON & LIONE, P.C.

STREET: 2000 K St., N.W., Suite 200  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20006

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/076.677  
 FILING DATE: 12-May-1998  
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/275,526  
 FILING DATE: 15-JUL-1994

ATTORNEY/AGENT INFORMATION:  
 NAME: Gadiano, Wilhem F.  
 REGISTRATION NUMBER: 37,136  
 REFERENCE/DOCKET NUMBER: 4121-49

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 429-0625  
 TELEFAX: (202) 293-0625  
 TELEX: 650 383 5605

INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1022 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: genomic DNA  
 ORIGINAL SOURCE:  
 ORGANISM: Bacillus pumilus  
 STRAIN: PRL B12  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-076-677-1

Query Match 36.7%; Score 273; DB 4; Length 1022;  
 Best Local Similarity 63.4%; Pred. No. 2.4e-75;  
 Matches 452; Conservative 0; Mismatches 255; Indels 6; Gaps 2;

QY 1 ATGAGACAAAAGAAATTCAGCTTGATTTAGCTTTTATGTTTGTGCTACTAACCTTA 60  
 DB 186 ATGAATTTGAAGAAGATTGAGCTGTGTTGTGATGTATGATTTGTCTGACACTG 245

QY 61 CCTCAGAAATAATTCAGGCAAAATGCTACCGACAATTCATTCGCAACACAGATGGC 120  
 DB 246 ACGGCTGTGCGGCTCATGCGAAACGATTTATGATATAGGATAGGACACACAGCGGA 305

QY 121 TATGATATGATTTTGAAGAGATAGCGTGGCTCTGGACAATGATTTCTCAATCAGGC 180  
 DB 306 TACGATTTGAAATTTATGGAAGATTACGGAATACCTCG---ATGACACTCAATAACGCG 362

QY 181 GGTACGTTCACTGCGCAATGGAACAATGTTTAAACAATATTTATTCGTAAGGTAAGAAA 240  
 DB 363 GGGCAATTTAGTCAAGCTGGAACAATATTGGAATGCTTTATTCGAAAAGGAAGAAG 422

QY 241 TTCAATGAACACAAACACACCAATGTTGGTAACATGTCCTAATAACTACGAGCCAAC 300  
 DB 423 TTTGATTCACATAAACTCATCACTTGGCAACATCTCCATCAACTCAACGACGCC 482

QY 301 TTCCAAACCAATGTAATGCGTATTTATGCGTCTATGTTGGAATGTTGACCTCTTGTG 360  
 DB 483 TTTAACCGGCGGGAATTCCTATTTATGTTGCTATGCTGGACACATCTCCATAGCT 542

QY 361 GAATATTATATTTGTCAGAGTTGGGCAACTGGCGTCCACAGGAGCAACGCCCTAAGGGG 420  
 DB 543 GAATATTATATTTGAGTTCATGGGCACTATCTGTCACAGG---AACGTATTAAGGA 599

QY 421 ACCATCACTGTTGATGGAGGACATATGATATCTACGAGACTCTTAGAGCTCAATCAACC 480  
 DB 600 TCATTTTATGCGGATGGAGGACATATGACATATATGAACGCTCCGCTGTCATCAGCCT 659

Query Match 36.7%; Score 273; DB 4; Length 1022;

NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAM  
STREET: 2000 K St.  
CITY: Washington

CITY: Wash



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 21:26:43 ; Search time 475.684 Seconds  
(without alignments)  
8421.106 Million cell updates/sec

Title: US-09-909-207-4  
Perfect score: 744  
Sequence: 1 ATGAGACAAAGAAATTCAC.....TAACCTTGGATATAAAACAAT 744

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3611042 seqs, 2692057975 residues

Total number of hits satisfying chosen parameters: 7222084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	744	100.0	744	US-09-909-207-4	Sequence 4, Appli
2	744	100.0	744	US-09-909-207-5	Sequence 5, Appli
3	744	100.0	1513	US-09-909-207-10	Sequence 10, Appl
4	744	100.0	1513	US-09-909-207-11	Sequence 11, Appl
5	663	89.1	663	US-09-909-207-1	Sequence 1, Appli
6	663	89.1	663	US-09-909-207-2	Sequence 2, Appli
7	119.4	16.0	1375	US-09-770-621-1	Sequence 1, Appli
8	119.4	16.0	1375	US-10-286-993-1	Sequence 1, Appli
9	115.2	15.5	596	US-10-307-441-39	Sequence 39, Appl
10	111.4	15.0	942	US-10-213-990-71	Sequence 71, Appl
11	110	14.8	1002	US-10-213-990-70	Sequence 70, Appl
12	108.8	14.6	705	US-10-213-990-68	Sequence 68, Appl

13	106	14.2	2898	15	US-10-299-393-1	Sequence 1, Appli
14	104.8	14.1	712	14	US-10-213-990-64	Sequence 64, Appl
15	103.6	13.9	666	14	US-10-213-990-65	Sequence 65, Appl
16	96.8	13.0	739	14	US-10-213-980-67	Sequence 67, Appl
17	86.8	11.7	678	10	US-09-803-454-3	Sequence 3, Appli
18	84.2	11.3	1023	16	US-10-244-596-6	Sequence 6, Appli
19	84	11.3	1027	16	US-10-244-596-2	Sequence 2, Appli
20	83.2	11.2	645	15	US-10-237-386-10	Sequence 10, Appl
21	83.2	11.2	657	15	US-10-237-386-11	Sequence 11, Appl
22	82.6	11.1	1011	16	US-10-244-596-11	Sequence 11, Appl
23	82.6	11.1	1011	16	US-10-244-596-12	Sequence 12, Appl
24	82.6	11.1	1011	16	US-10-244-596-13	Sequence 13, Appl
25	82.6	11.1	1023	16	US-10-244-596-5	Sequence 5, Appli
26	82.6	11.1	1023	16	US-10-244-596-7	Sequence 7, Appli
27	82.6	11.1	1023	16	US-10-244-596-8	Sequence 8, Appli
28	82.6	11.1	1023	16	US-10-244-596-9	Sequence 9, Appli
29	81	10.9	81	9	US-09-909-207-7	Sequence 7, Appli
30	81	10.9	81	9	US-09-909-207-8	Sequence 8, Appli
31	81	10.9	1023	16	US-10-244-596-3	Sequence 3, Appli
32	81	10.9	1023	16	US-10-244-596-10	Sequence 10, Appl
33	79.4	10.7	1023	16	US-10-244-596-4	Sequence 4, Appli
34	79	10.6	2225	10	US-03-790-070A-8	Sequence 8, Appli
35	77.8	10.5	1014	16	US-10-244-596-1	Sequence 1, Appli
36	77.2	10.4	749	18	US-10-425-115-82922	Sequence 82922, A
37	75.8	10.2	818	18	US-10-425-115-37173	Sequence 37173, A
38	73.8	9.9	850	18	US-10-425-115-177283	Sequence 177283, A
39	72.4	9.7	588	15	US-10-237-386-9	Sequence 9, Appli
40	72.4	9.7	983	9	US-09-467-368-1	Sequence 1, Appli
41	54.2	7.3	2054	15	US-10-419-969-5	Sequence 5, Appli
42	50.4	6.8	643	15	US-10-340-860A-39	Sequence 39, Appl
43	41	5.5	74	15	US-10-307-441-24	Sequence 24, Appl
44	40	5.4	557	9	US-09-970-616-1	Sequence 1, Appli
45	40	5.4	642	11	US-09-149-310-31	Sequence 31, Appl

## ALIGNMENTS

RESULT 1  
US-09-909-207-4  
; Sequence 4, Application US/09090207  
; Patent No. US20020115181A1

GENERAL INFORMATION:  
APPLICANT: ANDREE LAHAYE  
ERIC DE BUYL  
PIERRE LEDOUX  
RENE DETROZ

TITLE OF INVENTION: Xylanase, microorganisms produced it,  
DNA molecule, processes for preparation of this xylanase  
and uses thereof

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
STREET: 2000 K St., N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/09/909,207

APPLICATION NUMBER: 19-Jul-2001

FILING DATE: 19-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/470,953

FILING DATE: 06-JUNE-1995

ATTORNEY/AGENT INFORMATION:

NAME: Wilhem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136

REFERENCE/DOCKET NUMBER: 4121-40  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 744 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: Bacillus  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-909-207-4

Query Match 100.0%; Score 744; DB 9; Length 744;  
Best Local Similarity 100.0%; Pred. No. 2.1e-208;  
Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACAAAAGAAATTCACGCTGATTTAGCTTTTACCTTTTGTGTTTGTGCACTAACCTTA 60  
DB 1 ATGACACAAAAGAAATTCACGCTGATTTAGCTTTTACCTTTTGTGTTTGTGCACTAACCTTA 60

QY 61 CCTGCAGAAATAATTCAGGCAAAATCGTCCGACAAATTCATTCGCAACCCAGATGGC 120  
DB 61 CCTGCAGAAATAATTCAGGCAAAATCGTCCGACAAATTCATTCGCAACCCAGATGGC 120

QY 121 TATGATTAATGATTTTGAAGATAGCGTGGCTCTGGGACAAATGATTCATCAATCATGGC 180  
DB 121 TATGATTAATGATTTTGAAGATAGCGTGGCTCTGGGACAAATGATTCATCAATCATGGC 180

QY 181 GGTACGTTTCAGTCCCAATGGAACAAATGTTAAACATATATTATTCGTTAAAGGTAAATAA 240  
DB 181 GGTACGTTTCAGTCCCAATGGAACAAATGTTAAACATATATTATTCGTTAAAGGTAAATAA 240

QY 241 TTCAATGAAACAAACACACCAACCAAGTTGGTAAATGTCATTAATACCGAGCCCAAC 300  
DB 241 TTCAATGAAACAAACACACCAACCAAGTTGGTAAATGTCATTAATACCGAGCCCAAC 300

QY 301 TTCCCAACCAATGTAATGCGTATTTATGCGTCTATGTTGACCTGTTGACCTCTTTGTC 360  
DB 301 TTCCCAACCAATGTAATGCGTATTTATGCGTCTATGTTGACCTGTTGACCTCTTTGTC 360

QY 361 GAATATTAATGTCACAGTTGGGGCAACTGGCGTCCACAGGAGCAACCGCTTAAGGGG 420  
DB 361 GAATATTAATGTCACAGTTGGGGCAACTGGCGTCCACAGGAGCAACCGCTTAAGGGG 420

QY 421 ACCATCACTGTTGATGGAGGACATATGATCTACGAGACTCTTAGAGTCAATCAACCC 480  
DB 421 ACCATCACTGTTGATGGAGGACATATGATCTACGAGACTCTTAGAGTCAATCAACCC 480

QY 481 TCCATTAAAGGGATTGCCACATTTAAACAAATATGAGGTGTTTCGAAGATCGAAACGACG 540  
DB 481 TCCATTAAAGGGATTGCCACATTTAAACAAATATGAGGTGTTTCGAAGATCGAAACGACG 540

QY 541 AGTGGCAGATTTCTGTACGACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATG 600  
DB 541 AGTGGCAGATTTCTGTACGACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATG 600

QY 601 GGGAAATGATGAAGTCCGCTTACTGTAGAGGCTATCAAGTAGCGGAAGTGTAAAT 660  
DB 601 GGGAAATGATGAAGTCCGCTTACTGTAGAGGCTATCAAGTAGCGGAAGTGTAAAT 660

QY 661 GTATATAGCAATACATAAGAAATAACGGTAACCCCTCTCTCAATTAATGATGACGAG 720  
DB 661 GTATATAGCAATACATAAGAAATAACGGTAACCCCTCTCTCAATTAATGATGACGAG 720

QY 721 AGCAATACTTTGGATTAACCAAT 744  
DB 721 AGCAATACTTTGGATTAACCAAT 744

RESULT 2  
US-09-909-207-5  
; Sequence 5, Application US/09909207  
; Patent No. US20020115181A1  
; GENERAL INFORMATION:  
; APPLICANT: ANDREE LAHAYE  
; ERIC DE BUYL  
; PIERRE LEDOUX  
; RENE DETROZ  
; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
; DNA molecule, processes for preparation of this xylanase  
; and uses thereof  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WILLIAM BRINKS HOPFER GILSON & LIONE  
; STREET: 2000 K St., N.W., Suite 200  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/909,207  
; FILING DATE: 19-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/470,953  
; FILING DATE: 06-JUNE-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wilhem F. Gadiano, Esq.  
; REGISTRATION NUMBER: 37,136  
; REFERENCE/DOCKET NUMBER: 4121-40  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-429-0625  
; TELEFAX: (202) 293-1850  
; TELEX: 650 383-5605  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 744 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; STRAIN: Bacillus  
; NAME/KEY: CDS  
; LOCATION: 1..744  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: 82..744  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: 1..81  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-909-207-5

Query Match 100.0%; Score 744; DB 9; Length 744;  
Best Local Similarity 100.0%; Pred. No. 2.1e-208;  
Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACAAAAGAAATTCACGCTGATTTAGCTTTTACCTTTTGTGTTTGTGCACTAACCTTA 60  
DB 1 ATGACACAAAAGAAATTCACGCTGATTTAGCTTTTACCTTTTGTGTTTGTGCACTAACCTTA 60

QY 61 CCTGCAGAAATAATTCAGGCAAAATCGTCCGACAAATTCATTCGCAACCCAGATGGC 120  
DB 61 CCTGCAGAAATAATTCAGGCAAAATCGTCCGACAAATTCATTCGCAACCCAGATGGC 120



121 TATGATTATGAATTTGGAAAGATAGCGGTGGCTCTGGACAAATGATTTCTCAATCATGCG 180  
121 TATGATTATGAATTTGGAAAGATAGCGGTGGCTCTGGACAAATGATTTCTCAATCATGCG 180  
181 GGTACGTTTCAGTGGCCCAATGGAACAATGTTAAACAATATTTCCGTAAGGTAAGGTAAGG 240  
181 GGTACGTTTCAGTGGCCCAATGGAACAATGTTAAACAATATTTCCGTAAGGTAAGGTAAGG 240  
241 TTCAATGAAACACAAACACCAACCAAGTTGGTAACATGTCATTAACACTACGGAGCCAC 300  
241 TTCAATGAAACACAAACACCAACCAAGTTGGTAACATGTCATTAACACTACGGAGCCAC 300  
301 TTCCACCAAAATGGAATCGGTATTTATCGCTATGTTGACTGTTTCCACCTCTTGTG 360  
301 TTCCACCAAAATGGAATCGGTATTTATCGCTATGTTGACTGTTTCCACCTCTTGTG 360  
361 GAATATTATTTGTCGACAGTTGGGCAACTGGCGTCCACAGGAGCAACGGCTAAGGGG 420  
361 GAATATTATTTGTCGACAGTTGGGCAACTGGCGTCCACAGGAGCAACGGCTAAGGGG 420  
421 ACCATCACTGTTGATGGAGGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480  
421 ACCATCACTGTTGATGGAGGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480  
481 TCCATTAAAGGGGATTCACCAATTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGCG 540  
481 TCCATTAAAGGGGATTCACCAATTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGCG 540  
541 AGTGCCACGATTTCTGTCAGCAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 600  
541 AGTGCCACGATTTCTGTCAGCAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 600  
601 GGGAAATATGTAAGTTCGCGTACTGTAGAGGCTATCAAGTAGCGGAAGTGCTAAT 660  
601 GGGAAATATGTAAGTTCGCGTACTGTAGAGGCTATCAAGTAGCGGAAGTGCTAAT 660  
661 GTATATGACCAATACATTAAGATTAACGGTAAACCTCTCTCACTATTAGTATGACGAG 720  
661 GTATATGACCAATACATTAAGATTAACGGTAAACCTCTCTCACTATTAGTATGACGAG 720  
721 AGCAATACCTTTGGATAAAACCAAT 744  
721 AGCAATACCTTTGGATAAAACCAAT 744

RESULT 3  
US-09-909-207-10  
Sequence 10, Application US/09909207  
Patent No. US20020115181A1  
GENERAL INFORMATION:  
APPLICANT: ANDREE LAHAYE  
ERIC DE BUYL  
PIERRE LEDOUX  
RENE DETROZ  
TITLE OF INVENTION: Xylanase, microorganisms produced it,  
DNA molecule, processes for preparation of this xylanase  
and uses thereof  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAN BRINKS HOPER GILSON & LIONE  
STREET: 2000 K St., N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/909,207

FILING DATE: 19-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/470,953  
FILING DATE: 06-JUNE-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilhem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-40  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-0825  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1513 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: Bacillus  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-09-909-207-10

Query Match 100.0%; Score 744; DB 9; Length 1513;  
Best Local Similarity 100.0%; Pred. No. 3e-208;  
Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAGACAAAAGAAATGACGTTGATTTTGTAGCTTTTGTAGTTTGTGCTTAACTTA 60  
DB 620 ATGAGACAAAAGAAATGACGTTGATTTTGTAGCTTTTGTAGTTTGTGCTTAACTTA 679  
QY 61 CTGCGAAGAAATTTTCAGGCACAAATCGTCACGCAATTCCTTGGACCAACGATGCG 120  
DB 680 CTGCGAAGAAATTTTCAGGCACAAATCGTCACGCAATTCCTTGGACCAACGATGCG 739  
QY 121 TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTGGGACAAATGATTTCTCAATCATGCG 180  
DB 740 TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTGGGACAAATGATTTCTCAATCATGCG 799  
QY 181 GGTACGTTTCAGTGGCCCAATGGAACAATGTTTAAACAATATTTCCGTAAGGTAAGGTAAGG 240  
DB 800 GGTACGTTTCAGTGGCCCAATGGAACAATGTTTAAACAATATTTCCGTAAGGTAAGGTAAGG 859  
QY 241 TTCCAAACCAATGTTTAAACAATGGAACAATGTTTAAACAATGTTTAAACAATGTTTAAACAAT 300  
DB 860 TTCCAAACCAATGTTTAAACAATGGAACAATGTTTAAACAATGTTTAAACAATGTTTAAACAAT 919  
QY 301 TTCCAAACCAATGTTTAAACAATGGAACAATGTTTAAACAATGTTTAAACAATGTTTAAACAAT 360  
DB 920 TTCCAAACCAATGTTTAAACAATGGAACAATGTTTAAACAATGTTTAAACAATGTTTAAACAAT 979  
QY 361 GAATATTATTTGTCGACAGTTTGGGCAACTGGCGTCCACAGGAGCAACGGCTAAGGGG 420  
DB 980 GAATATTATTTGTCGACAGTTTGGGCAACTGGCGTCCACAGGAGCAACGGCTAAGGGG 1039  
QY 421 ACCATCACTGTTGATGGAGGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480  
DB 1040 ACCATCACTGTTGATGGAGGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 1099  
QY 481 TCCATTAAAGGGGATTCGCCCAATTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGCG 540  
DB 1100 TCCATTAAAGGGGATTCGCCCAATTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGCG 1159  
QY 541 AGTGCCACGATTTCTCTCAGCAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 600  
DB 1160 AGTGCCACGATTTCTCTCAGCAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG 1219  
QY 601 GGGAAATATGTAAGTTCGCGTACTGTAGAGGCTATCAAGTAGCGGAAGTGCTAAT 660  
DB 1220 GGGAAATATGTAAGTTCGCGTACTGTAGAGGCTATCAAGTAGCGGAAGTGCTAAT 1279

QY 661 GTATATAGCAATACACTAAGAAATTAAACGGTAACCCCTCTCTCAACTATTAGTAATGACGAG 720  
 DB 1280 GTATATAGCAATACACTAAGAAATTAAACGGTAACCCCTCTCTCAACTATTAGTAATGACGAG 1339  
 QY 721 AGCATAACTTTGGATAAAAAACAAT 744  
 DB 1340 AGCATAACTTTGGATAAAAAACAAT 1363

RESULT 4

US-09-909-207-11  
 ; Sequence 11, Application US/09909207  
 ; Patent No. US20020115181A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ANDREE LAHAYE  
 ; ERIC DE BUYL  
 ; PIERRE LEDOUX  
 ; RENE DETROZ  
 ; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
 ; DNA molecule, processes for preparation of this xylanase  
 ; and uses thereof  
 ; NUMBER OF SEQUENCES: 29  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
 ; STREET: 2000 K St., N.W., Suite 200  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/909,207  
 FILING DATE: 19-Jul-2001  
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/470,953  
 FILING DATE: 06-JUNE-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Wilhem F. Gadiano, Esq.  
 REGISTRATION NUMBER: 37,136  
 REFERENCE/DOCKET NUMBER: 4121-40

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-429-0625  
 TELEFAX: (202) 293-1850  
 TELEX: 650 383-5605

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 1513 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 ORIGINAL SOURCE:  
 STRAIN: Bacillus

FEATURE:

NAME/KEY: CDS  
 LOCATION: 620..1363  
 FEATURE:  
 NAME/KEY: mat\_peptide  
 LOCATION: 701..1363  
 FEATURE:  
 NAME/KEY: sig\_peptide  
 LOCATION: 620..700

SEQUENCE DESCRIPTION: SEQ ID NO: 11:

US-09-909-207-11  
 Query Match 100.0%; Score 744; DB 9; Length 1513;  
 Best Local Similarity 100.0%; Pred. No. 3e-208;  
 Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGACAAAAGAAATTGACGTTGATTTTATAGCCCTTTTATGTTTGTGTTTGCACCTAACCTTA 60  
 DB 620 ATGAGACAAAAGAAATTGACGTTGATTTTATAGCCCTTTTATGTTTGTGTTTGCACCTAACCTTA 679  
 QY 61 CCTGCAGAAAATAATTTCAGGCACAAAATCGTCACCGACAAATTCCTCAATTTGGCAACACGATGGC 120  
 DB 680 CCTGCAGAAAATAATTTCAGGCACAAAATCGTCACCGACAAATTCCTCAATTTGGCAACACGATGGC 739  
 QY 121 TATGATTATGAATTTTGGAAAAGATAGCGGTGGCTCTGGGACAATGATTTCTCAATCATGGC 180  
 DB 740 TATGATTATGAATTTTGGAAAAGATAGCGGTGGCTCTGGGACAATGATTTCTCAATCATGGC 799  
 QY 181 GGTACGTTTCAAGTCCCAATGGAACAATGTTAAACAATATTTATTCCTGTAAGAGTAAAGAAA 240  
 DB 800 GGTACGTTTCAAGTCCCAATGGAACAATGTTAAACAATATTTATTCCTGTAAGAGTAAAGAAA 859  
 QY 241 TTCAATGAAAACACAAACACACCAACCAAGTTGGTAACATGTCATATAAATACTACGGAGCCAAC 300  
 DB 860 TTCAATGAAAACACAAACACACCAACCAAGTTGGTAACATGTCATATAAATACTACGGAGCCAAC 919  
 QY 301 TTCCAACCAATGGTAATGCGTATTTATGCGTCTATGTTGGTGGACTGTGACCCCTCTTGTG 360  
 DB 920 TTCCAACCAATGGTAATGCGTATTTATGCGTCTATGTTGGTGGACTGTGACCCCTCTTGTG 979  
 QY 361 GATATATTATTTGTCGACAGTTGGGCAACTGGCGTCCACGAGGACCAACGCTTAAGGGG 420  
 DB 980 GATATATTATTTGTCGACAGTTGGGCAACTGGCGTCCACGAGGACCAACGCTTAAGGGG 1039  
 QY 421 ACCATCACTGTTGATGGAGGAACATATGATATCTACGAGACTCTTTAGAGTCAATCAACCC 480  
 DB 1040 ACCATCACTGTTGATGGAGGAACATATGATATCTACGAGACTCTTTAGAGTCAATCAACCC 1099  
 QY 481 TCCATTAAGGGGATGGCACAATTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACG 540  
 DB 1100 TCCATTAAGGGGATGGCACAATTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACG 1159  
 QY 541 AGTGGCAGATTTCTGTCAGCAACCACTTTAGAGCGTGGGAAAACCTTAGGATGAATATG 600  
 DB 1160 AGTGGCAGATTTCTGTCAGCAACCACTTTAGAGCGTGGGAAAACCTTAGGATGAATATG 1219  
 QY 601 GGGAAAATGTATGAAGTCGGCTTACTGTAGAAAGCTATCAAGTAGCGGAAGTGTAAAT 660  
 DB 1220 GGGAAAATGTATGAAGTCGGCTTACTGTAGAAAGCTATCAAGTAGCGGAAGTGTAAAT 1279  
 QY 561 GTATATAGCAATACACTAAGAAATTAAACGGTAACCCCTCTCTCAACTATTAGTAATGACGAG 720  
 DB 1280 GTATATAGCAATACACTAAGAAATTAAACGGTAACCCCTCTCTCAACTATTAGTAATGACGAG 1339  
 QY 721 AGCATAACTTTGGATAAAAAACAAT 744  
 DB 1340 AGCATAACTTTGGATAAAAAACAAT 1363

RESULT 5

US-09-909-207-1  
 ; Sequence 1, Application US/09909207  
 ; Patent No. US20020115181A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ANDREE LAHAYE  
 ; ERIC DE BUYL  
 ; PIERRE LEDOUX  
 ; RENE DETROZ

TITLE OF INVENTION: Xylanase, microorganisms produced it,

DNA molecule, processes for preparation of this xylanase  
 and uses thereof

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
 STREET: 2000 K St., N.W., Suite 200  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.

ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/909,207  
FILING DATE: 19-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/470,953  
FILING DATE: 06-JUNE-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilhem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-40  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 663 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: Bacillus  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-909-207-1  
Query Match 89.1%; Score 663; DB 9; Length 663;  
Best Local Similarity 100.0%; Pred. No. 1.4e-184;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 82 CAAATCGTCACGACAAATTCATTGGCAACCCAGATGGCTATGATTAATGAAATTTGGAAA 141  
DB 1 CAAATCGTCACGACAAATTCATTGGCAACCCAGATGGCTATGATTAATGAAATTTGGAAA 60  
QY 142 GATAGCGGTGGCTCTGGGCAATGATTCATGATGGGGTACGTTTCAGTGCCTCAATGG 201  
DB 61 GATAGCGGTGGCTCTGGGCAATGATTCATGATGGGGTACGTTTCAGTGCCTCAATGG 120  
QY 202 AACATGTTAAACAATATATTTCGTAAAGGTAAATAATTCATGAAGAACACAAACACAC 261  
DB 121 AACATGTTAAACAATATATTTCGTAAAGGTAAATAATTCATGAAGAACACAAACACAC 180  
QY 262 CAACAAGTTGGTAACATGTCATATAACTAGGAGCCAACTTCCAAACCAATGTAATGG 321  
DB 181 CAACAAGTTGGTAACATGTCATATAACTAGGAGCCAACTTCCAAACCAATGTAATGG 240  
QY 322 TATTTATGCTCTATGTTGACTGTGACCTCTTGTCGAATATTATATTGTCGACAGT 381  
DB 241 TATTTATGCTCTATGTTGACTGTGACCTCTTGTCGAATATTATATTGTCGACAGT 300  
QY 382 TGGGGCAACTGGCGTCCACAGAGCAACCCCTAAAGGGGACCATCATCTGTTGATGGAGGA 441  
DB 301 TGGGGCAACTGGCGTCCACAGAGCAACCCCTAAAGGGGACCATCATCTGTTGATGGAGGA 360  
QY 442 ACATATGATATCTACGAGATCTTAGAGTCAATCAACCTCCATTAAAGGGGATGGCCACA 501  
DB 361 ACATATGATATCTACGAGATCTTAGAGTCAATCAACCTCCATTAAAGGGGATGGCCACA 420  
QY 502 TTTTAAACAATTTGGAGTGTTCGAAGATCAAAACGACGACGATGGCAGATTTCTGTGACG 561  
DB 421 TTTTAAACAATTTGGAGTGTTCGAAGATCAAAACGACGACGATGGCAGATTTCTGTGACG 480  
QY 562 AACCACTTTAGAGCGTGGGAAAACTTTAGGGATGAATATGGGGAATATGATGAAGTCGCG 621  
DB 481 AACCACTTTAGAGCGTGGGAAAACTTTAGGGATGAATATGGGGAATATGATGAAGTCGCG 540

QY 622 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTCTAATGTATATAGCAATACACTAAGA 681  
DB 541 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTCTAATGTATATAGCAATACACTAAGA 600  
QY 682 ATTAACGGTAACCTCTCTCAACTATTAGTATGACGAGAGCATAACTTTGGATAAAAC 741  
DB 601 ATTAACGGTAACCTCTCTCAACTATTAGTATGACGAGAGCATAACTTTGGATAAAAC 660  
QY 742 AAT 744  
DB 661 AAT 663  
RESULT 6  
US-09-909-207-2  
Sequence 2, Application US/09909207  
Patent No. US20020115181A1  
GENERAL INFORMATION:  
APPLICANT: ANDREE LAHAYE  
ERIC DE BUYL  
PIERRE LEDOUX  
RENE DETROZ  
TITLE OF INVENTION: Xylanase, microorganisms produced it,  
DNA molecule, processes for preparation of this xylanase  
and uses thereof  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAN BRINKS HOPER GILSON & LIONE  
STREET: 2000 K St., N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/909,207  
FILING DATE: 19-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/470,953  
FILING DATE: 06-JUNE-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilhem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-40  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 663 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: Bacillus  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 1..663  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..663  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-909-207-2  
Query Match 89.1%; Score 663; DB 9; Length 663;  
Best Local Similarity 100.0%; Pred. No. 1.4e-184;

Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 CAAATCGTCACGACAAATTCATTCGGCAACACGATGCTATGATTAATTTGGAAA 141  
 Db 1 CAAATCGTCACGACAAATTCATTCGGCAACACGATGCTATGATTAATTTGGAAA 60

QY 142 GATAGCGGTGCTCTGGGACAAATGATTCATCATGCGGTGACGTTCAAGTCCCAATGG 201  
 Db 61 GATAGCGGTGCTCTGGGACAAATGATTCATCATGCGGTGACGTTCAAGTCCCAATGG 120

QY 202 AACATGTTAACACATATATTCGTAAGGTAAATTCATGAAACCAACACAC 261  
 Db 121 AACATGTTAACACATATATTCGTAAGGTAAATTCATGAAACCAACACAC 180

QY 262 CAACAAGTTGTAACATGTCATTAACATGACGAGCAACCTCCACCAAAATGTAATCGG 321  
 Db 181 CAACAAGTTGTAACATGTCATTAACATGACGAGCAACCTCCACCAAAATGTAATCGG 240

QY 322 TATTTATGCGTCTATGTTGGACTGTTGACCCCTCTTGTGCAATATTAATTTGCGACAGT 381  
 Db 241 TATTTATGCGTCTATGTTGGACTGTTGACCCCTCTTGTGCAATATTAATTTGCGACAGT 300

QY 382 TGGGCAACTGCGTCCACGAGGACCAAGCTTAAGGGACCATCACTGTTGATGGAGGA 441  
 Db 301 TGGGCAACTGCGTCCACGAGGACCAAGCTTAAGGGACCATCACTGTTGATGGAGGA 360

QY 442 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCTATTAAAGGGATTGCCACA 501  
 Db 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCTATTAAAGGGATTGCCACA 420

QY 502 TTAAACAATATTTGAGTGTTCGAAGATCGAAACGACAGTGGCAGCAATTTCTGTACGC 561  
 Db 421 TTAAACAATATTTGAGTGTTCGAAGATCGAAACGACAGTGGCAGCAATTTCTGTACGC 480

QY 562 RACCACTTTAGAGCTGGGAAACTTAGGATGAATATGGGAAATGTATGAAGTCGC 621  
 Db 481 RACCACTTTAGAGCTGGGAAACTTAGGATGAATATGGGAAATGTATGAAGTCGC 540

QY 622 CTTACTGTAGAGGCTATCAAGTACGGAAGTGTCTATGATATAGCAATACACTAGA 681  
 Db 541 CTTACTGTAGAGGCTATCAAGTACGGAAGTGTCTATGATATAGCAATACACTAGA 600

QY 682 ATTACCGTAACCTCTCTCACTATTATGATGACGAGACATACTTTGATAAAAC 741  
 Db 601 ATTACCGTAACCTCTCTCACTATTATGATGACGAGACATACTTTGATAAAAC 660

QY 742 AAT 744  
 Db 661 AAT 663

RESULT 7

US-09-770-621-1  
 ; Sequence 1, Application US/09770621  
 ; Patent No. US20010024815A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: M ntyl, Arja  
 ; APPLICANT: Vehmaanper, Jari  
 ; APPLICANT: Fagerstr m, Richard  
 ; APPLICANT: Lantto, Raija  
 ; APPLICANT: Paloheimo, Marja  
 ; APPLICANT: Suominen, Pirkko  
 ; APPLICANT: Lahtinen, Tarja  
 ; TITLE OF INVENTION: Production and Secretion of Proteins of  
 ; NUMBER OF SEQUENCES: 39  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.  
 ; STREET: 1100 New York Ave., N.W. Suite 600  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20005  
 ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/770,621  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/590,563  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/332,412  
 ; FILING DATE: 31-OCT-1994  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/282,001  
 ; FILING DATE: 29-JUL-1994  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Bugaisky, Lawrence B.  
 ; REGISTRATION NUMBER: 35,086  
 ; REFERENCE/DOCKET NUMBER: 1050.0340003  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-371-2600  
 ; TELEFAX: 202-371-2540  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1375 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: No. US20010024815A1 Relevant  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 303..1334  
 ; US-09-770-621-1

Query Match 16.0%; Score 119.4; DB 9; Length 1375;

Best Local Similarity 57.1%; Pred. No. 3.1e-24; Mismatches 176; Indels 3; Gaps 1;

Matches 238; Conservative 0; Mismatches 176; Indels 3; Gaps 1;

QY 279 GTCCATAAACTAGCGAGCCCAACTTCCAAACCAATGTAATCGGTATTTATCGGTCTATGG 338  
 Db 608 GACCGTGACCTACAAACGCTCTTCAACCCGTCGGGTAAACGGCTACCTCAAGCTCTACGG 667

QY 339 TTGAGCTGTTGACCCCTCTTGTGCAATATTAATTTGTCGACAGTTGGGGCAACTGGCGTCC 398  
 Db 668 CTGGACCAGGAACCCGCTCGTCGAGTACTACATCGTCGAGAGCTGGGGCACCTACCGGCC 727

QY 399 ACCAGGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGAACATATGATATCTACGA 458  
 Db 728 CACCGG---CACCTACAAAGGGACCGGTCAACACGCGGGGACGTAACGACATCTACGA 784

QY 459 GACTCTTTAGAGTCAATCAACCCCTCCATTAAAGGGGATTCGCCACATTTAAACAATATTGGAG 518  
 Db 785 GACTGCGGTGACAAACGCGCGCTCCATCGAGGGCACCCGACCTTCCAGCAGTTCTGGAG 844

QY 519 TGTTGGAAGATCGAAACGCAAGTGGCAGATTTCTGTGTCAGCAACCACTTTAGAGCGTG 578  
 Db 845 CGTCCGGCAGCAGAAGCGGACAGCGGCACCATCAACATCGGCAACCACTTCGACGCGCTG 904

QY 579 GGAATACTTAGGGATGAATATGGGAAATGTAAGTCGCTTACTCTAGAGGCTA 638  
 Db 905 GGCCCGCGCGCATGAACCTGGGACGACGACTACCATCATGCGGACCGAGGCTA 964

QY 639 TCAAAGTAGCGGAGTGTCTAATGTATATAGCAATATACACTAAGAAATTAACGGTAACCC 695  
 Db 965 CCAGAGCAGCGGTAGTCTCCACCGTCTCCATCAGCGGGGTGGCAACCCCGGCAACCC 1021

RESULT 8

```
US-10-286-993-1
; Sequence 1, Application US/10286993
; Publication No. US20030148453A1
; GENERAL INFORMATION:
; APPLICANT: Mäntylä, Arja
; APPLICANT: Paloheimo, Marja
; APPLICANT: Lantto, Raija
; APPLICANT: Fagerstrom, Richard
; APPLICANT: Lahtinen, Tarja
; APPLICANT: Suominen, Pirkko
; APPLICANT: Vehmaampere, Jari
; TITLE OF INVENTION: Production and Secretion of Proteins in Filamentous
; TITLE OF INVENTION: Fungi
; FILE REFERENCE: 1716.0340004
; CURRENT APPLICATION NUMBER: US/10/286,993
; PRIOR FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/120,804
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: PCT/FI97/00037
; PRIOR FILING DATE: 1997-01-24
; PRIOR APPLICATION NUMBER: US 08/590,563
; PRIOR FILING DATE: 1996-01-26
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1375
; TYPE: DNA
; ORGANISM: Actinomadura flexuosa (Strain: DSM43186)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (303)...(1337)
; OTHER INFORMATION: Product= AM35 xylanase
; US-10-286-993-1

Query Match 16.0%; Score 119.4; DB 15; Length 1375;
Best Local Similarity 57.1%; Pred. No. 3.1e-24;
Matches 238; Conservative 0; Mismatches 176; Indels 3; Gaps 1;

QY 279 GTCCATPAATACTAGCGAGCCAACTTCCAAACAAATGTAATGCGTATTATGCGTCTATGG 338
DB 608 GACCGTGACCTACAACGCCCTCTTCAACCCGTCGGGTACGGCTACCTCAACGCTCTACGG 667
QY 339 TTGAGCTGTGACCTCTTGTGCAATATTATTTGTCGACAGTTGGGGCAACTGGCGTCC 398
DB 668 CTGGACAGCAACCCGCTGTCGAGTACTACATCGTCGAGAGCTGGGGCACTTACCGGCC 727
QY 399 ACCAGGAGCAACCCCTAAGGGGACCATCACTGTTGATGGAGGAACATATGATATCTACGA 458
DB 728 CACCGG---CACCTACAGGGGACCGTCAACCCGACGGGGAACTGACATCTACGA 784
QY 459 GACTCTTAGAGTCAATCAACCTCCATTAAAGGGGATTCGACATTTTAAACAATATTGGAG 518
DB 785 GACCTGGCGGTACAACCGCGCGTCCATCGAGGGCAACCCGACCTTCCAGCAGTTCTGGAG 844
QY 519 TGTTCGAAGATCGAAACGACGAGTGGCAGGATTTCTGTGACGAAACCACTTTAGACGCTG 578
DB 845 CGTCCGCGCAGCAGAAGCGGACGAGCGGCACCATCACCATCGGCAACCACTTGCACGCTG 904
QY 579 GGAACACTTAGGATGAATATGGGAAATGATGAAGTCGCGCTTACTGTAGAGGCTA 638
DB 905 GGCCCGCGCGCATGAACCTGGGCGCCACGACTACCATGATGGGACCGAGGGCTA 964
QY 639 TCAAAGTAGCGGAGTGCTAATGTATATATATACCAATACATTAAGAAATTAACGGTAACCC 695
DB 965 CCAGAGCAGCGGTAGTCTCCACCGTCTCCATCAGCGGGTGGCAACCCCGGCAACCC 1021

RESULT 9
US-10-307-441-39
; Sequence 39, Application US/10307441
; Publication No. US20030166236A1
; GENERAL INFORMATION:
; APPLICANT: SUNG, Wing L.

US-10-286-993-1
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Modified Xylanases Exhibiting Increased Thermophilicity
; FILE REFERENCE: 027367-5006US
; CURRENT APPLICATION NUMBER: US/10/307,441
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: PCT/CA01/00769
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/213,803
; PRIOR FILING DATE: 2000-05-31
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 596
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Trx
; US-10-307-441-39

Query Match 15.5%; Score 115.2; DB 15; Length 596;
Best Local Similarity 57.7%; Pred. No. 3.5e-23;
Matches 226; Conservative 0; Mismatches 163; Indels 3; Gaps 1;

QY 274 AACATGTCCATAAACTACGAGCCAACTTCCAAACCAAAATGGTAATGCGTATTATGCGTC 333
DB 191 AATAAGGTGATCAACTTCTCTGGATCTTATATCCGATGGGAATTCATCTTAAGCGTC 250
QY 334 TATGTTTGGACTGTTGACCCCTTTCGAAATATTATTTGTCGACAGTTGGGGCAACTGG 393
DB 251 TATGCTGGTCTAGAAACCCACTGATTGAATATTATTCATTTGCGAAATTTTCGGTACCTAC 310
QY 394 CGTCC---ACCAGGAGCAACGCTTAAGGGGACCATCACTGTTGATGGGAGGAACATATGAT 450
DB 311 AATCCGAGTACCGCGGCCCAAAATTTAGCGGAAGTCATAGTGGATCCGTATATGAT 370
QY 451 ATCTACGAGACTCTTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCCACATTTAAACAA 510
DB 371 ATCTACCGTACCACCGGTTTAAATCAGCATGATGATGGAACCGCCACCTTTATCAG 430
QY 511 TATTTGAGTGTTCGAAGATCGAAACGACGAGTGGCAGGATTTCTGTGACGAAACCACTTT 570
DB 431 TACTGGAGTGTAGACGTAATCATCGGAGCTCCGGTTCGGTTAATACTCGGAATCACTTT 490
QY 571 AGAGCGTGGGAAACTTAGGATGAATATGGGAAATGATATGAGTCCGCGCTTACTGTA 630
DB 491 AATGATGGGCACAGCAAGGGTTAACCCCTAGGTACAATGGATTATCAAAATCGTAGCGGTG 550
QY 631 GAAGGCTATCAAAAGTAGCGGAAGTGCCTAATGT 662
DB 551 GAAGGCTACTTCTCGAGTGGTTCGCGTAGTAT 582

RESULT 10
US-10-213-990-71
; Sequence 71, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Storms, Reg
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
; TITLE OF INVENTION: ENZYMES AND METHODS OF USE
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Aspergillus
```

```
FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(942)
US-10-213-990-71

Query Match      15.0%; Score 111.4; DB 14; Length 942;
Best Local Similarity 57.0%; Pred. No. 5.9e-22;
Matches 224; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Qy 263 AACAAAGTTGGTAACATGTCCTAAACTACGGAGCCAACTTCCAAACCAAAATGGTAATGGT 322
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 257 ATCCAGGGAGTGACCATGACATTAACCTTCTCTGGCAGCTTCAATCTCTCCGGAAATGCTT 316
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 323 ATTTATGCTCTATGGTTGGAGCTGTGACCTCTTGTGCAATATATATATTTGTACAGTT 382
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 317 ACCTGTCCGTGTATGGATGGAATACCAACCCCTAGTCGAATACTACATCTCCAGAACT 376
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 383 GGGCA---ACTGGCGTCCACAGGAGCAACGCTTAAGGGAGCACTACTGTTCATGGAG 439
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 377 ATGGCAGTTCAATCTCGCTCGGCATGACGCAAGGGCACCGTCAACGAGCATGGAT 436
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 440 GAACATATGATATACAGACTCTTTAGAGTCAATCAACCTTCCATTAAAGGGGATTCGA 499
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 437 CCACCTACGACATCTATGAGCACCACAAAGGTCAACAGCTTCGATCTCGGCACGGCA 496
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 500 CATTAAACATATTTGGAGTGTTCGAAGATCGAAACGACAGTGGCAGCATTTCTGTCA 559
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 497 CTTCAACCAATACTGGTCCATCGCCCAAAACAAAGCGATCCAGCGGCACAGTCAACCCG 556
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 560 GCAACCACTTTAGAGCGTGGGAAACTTTAGGATGAATATGGGAAATGATGAAGTCG 619
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 557 CGAATCACTTCAGGCTGGGCTAGTCTGGGGATGAACTGGGTACCCATFACTATCAGA 616
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 620 CGCTTACTGTAGAAGCTATCAAGTAGCGGAA 652
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 617 TTGTTTCCACTGAGGATATGAGACGCGGTA 649
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-10-213-990-70
; Sequence 70, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Storms, Reg
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 1002
; TYPE: DNA
; ORGANISM: Aspergillus
US-10-213-990-70

Query Match      14.8%; Score 110; DB 14; Length 1002;
Best Local Similarity 57.8%; Pred. No. 1.6e-21;
Matches 216; Conservative 0; Mismatches 155; Indels 3; Gaps 1;

Qy 282 CATAAATACGAGCCAACTTCCAAACCAATGGTAATGGTATTTATCGTCTATGGTTG 341
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 336 CATTACCTTCTCTGGCAGCTTCAATCTCTCGGAAATGCTTACCTGTCCGTGTATGGATG 395
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 342 GACTGTTGACCTCTTGTGCAATATATATGTCGACAGTTGGGCA---ACTGGCGTCC 398
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 396 GACTACCAACCCCTTAGTCGAATCTACATCTCCAGAACTATGCGCAATCAATCTCTGG 455
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 399 ACCAGGAGCAACCCCTTAAGGGGACCATCACTCTTTGATGGAGAACATATGATATCTACGA 458
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(942)
US-10-213-990-71

Query Match      15.0%; Score 111.4; DB 14; Length 942;
Best Local Similarity 57.0%; Pred. No. 5.9e-22;
Matches 224; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

Qy 263 AACAAAGTTGGTAACATGTCCTAAACTACGGAGCCAACTTCCAAACCAAAATGGTAATGGT 322
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 257 ATCCAGGGAGTGACCATGACATTAACCTTCTCTGGCAGCTTCAATCTCTCCGGAAATGCTT 316
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 323 ATTTATGCTCTATGGTTGGAGCTGTGACCTCTTGTGCAATATATATATTTGTACAGTT 382
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 317 ACCTGTCCGTGTATGGATGGAATACCAACCCCTAGTCGAATACTACATCTCCAGAACT 376
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 383 GGGCA---ACTGGCGTCCACAGGAGCAACGCTTAAGGGAGCACTACTGTTCATGGAG 439
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 377 ATGGCAGTTCAATCTCGCTCGGCATGACGCAAGGGCACCGTCAACGAGCATGGAT 436
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 440 GAACATATGATATACAGACTCTTTAGAGTCAATCAACCTTCCATTAAAGGGGATTCGA 499
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 437 CCACCTACGACATCTATGAGCACCACAAAGGTCAACAGCTTCGATCTCGGCACGGCA 496
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 500 CATTAAACATATTTGGAGTGTTCGAAGATCGAAACGACAGTGGCAGCATTTCTGTCA 559
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 497 CTTCAACCAATACTGGTCCATCGCCCAAAACAAAGCGATCCAGCGGCACAGTCAACCCG 556
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 560 GCAACCACTTTAGAGCGTGGGAAACTTTAGGATGAATATGGGAAATGATGAAGTCG 619
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 557 CGAATCACTTCAGGCTGGGCTAGTCTGGGGATGAACTGGGTACCCATFACTATCAGA 616
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 620 CGCTTACTGTAGAAGCTATCAAGTAGCGGAA 652
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 617 TTGTTTCCACTGAGGATATGAGACGCGGTA 649
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-10-213-990-68
; Sequence 68, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Storms, Reg
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Aspergillus
; NAME/KEY: CDS
; LOCATION: (1)...(705)
US-10-213-990-68

Query Match      14.6%; Score 108.8; DB 14; Length 705;
Best Local Similarity 51.9%; Pred. No. 3e-21;
Matches 298; Conservative 0; Mismatches 267; Indels 9; Gaps 2;

Qy 87 CGTCAACGACAAATTCCTATTCGCAACCAACGATGGCTATGATGATGAAATTTTGGAAAGATAG 146
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 120 CGGCACCCCAAGCTCCACCGCTGGAAACACGGCTACTACTCTCTCTGGACTGATGG 179
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 147 CGGTGGCTCTGGGACAATGATTTCAATCATATGGCGGTACGTTCACTAGTCCCAATGGAAACA 206
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 180 CGGCGGCGACGTGACCTACACCAATGGCGCGGTGGCTCGTACTCCGTCAACTGGAGGAA 239
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 207 TGTTAACACATATATTCGTTAAAGGTAAATAAATTCATGAACACACACACACACACACAC 266
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 240 CGTGGCAACTTTGTCTGGTGGAAAGGGC-----TGGAAACCTTGGAGCGGTAGGTACCG 293
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 267 AGTTGGTAACATGTCCTAAACTACGGAGCCAACTTCCAAACCAATGGTAATGGTATTT 326
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 294 AGCTTTAGTAGAACCATCACTACGGAGGAGCTTCAACCCCGGCGCATGGCTACTCT 353
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 327 ATGCGTCTATGGTTGGACTGTGTGACCTCTTGTGCAATATATATTTGTGCAAGTTGGGG 386
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 354 GGCTGTCTACGCTGGACCAACCAACCCCTTGATTCAGTACTACGTTGTTGAGTCTGTATG 413
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 387 CA---ACTGGCGTCCACGAGGACACGCTTAAGGGGACCATCACTGTGTTGATGGAGAAC 443
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 414 TACATACAAACCCCGGCGCGGTACTCTTCAGGGGCACTGTCTCAACCCGACGGTGGCAC 473
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

444 ATATGATATCTAGGAGACTCTTAGAGTCAATCAACCCCTCCATTAGGGGATGGCCACATT 503  
 474 TTACAAATCTACACGGCGTTTCGTAACAATGCTCCCTCCATCGAAGGACCAAGACCTT 533  
 504 TAAACAATATTGGAGTGTTCGAAGATCGAAACGACAGTGGCAGATTTCTGTGAGCAA 563  
 534 CACCCAGTACTGCTGTGCGGACCTCCAAAGGTACCGGGGCACTGTCAACCATGGCCAA 593  
 564 CCACCTTTAGAGCGTGGGAAAACCTTAGGGATGAATATGGGAAAATGTATGAAGTGGCGCT 623  
 594 CCACCTTCAACGCTGGAGCAGACTGGGCATGAACCTGGGAACTCACAACCTACAGATTGT 653  
 624 TACTGTAGAGGCTATCAAGTAGCGGAGTGTCT 657  
 654 CGCCACTGAGGGTTACGAGCAGCGGATCTGCT 687

RESULT 13  
 US-10-299-393-1  
 ; Sequence 1, Application US/10299393  
 ; Publication No. US20030108642A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sabatier, Alain  
 ; APPLICANT: Fish, Neville Marshall  
 ; APPLICANT: Haigh, Nigel Paterson  
 ; TITLE OF INVENTION: PENICILLIUM FUNICULOSUM STRAIN USEFUL  
 ; FILE REFERENCE: A32917-PCT-USA-1 (072667.0183)  
 ; CURRENT APPLICATION NUMBER: US/10/299,393  
 ; CURRENT FILING DATE: 2002-11-19  
 ; PRIOR APPLICATION NUMBER: 09/462,246  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: EPO 9801161.5  
 ; PRIOR FILING DATE: 1998-05-06  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 2898  
 ; TYPE: DNA  
 ; ORGANISM: Penicillium funiculosum  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1317)...(1589)  
 ; NAME/KEY: intron  
 ; LOCATION: (1590)...(1642)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1644)...(2042)  
 ; FEATURE:  
 ; NAME/KEY: RBS  
 ; LOCATION: (570)...(576)  
 ; FEATURE:  
 ; NAME/KEY: RBS  
 ; LOCATION: (724)...(730)  
 ; US-10-299-393-1

Query Match 14.2%; Score 106; DB 15; Length 2898;  
 Best Local Similarity 56.5%; Pred. No. 4.1e-20;  
 Matches 218; Conservative 0; Mismatches 165; Indels 3; Gaps 1;  
 276 CATGTCCATAAATACGAGGCCAACTTCCAAACAAATGGTAATTTATGCGTCTA 335  
 1640 CAGGACTGTACGTACTCAGGAGAAATTAACCCCTCTGGAACGCTTTATTTGGCTGTCTA 1699  
 336 TGGTTGACATGTTGACCCCTTTGCGAATATTATATGTCGACAGTTGGGGCACTGGCG 395  
 1700 CGGGTGACAACACAGATCCTCTTGTGGAATACTACATCCTGAGTCTTACGGCACCTATAA 1759  
 396 TCCACCA---GGAGCAACGCTTAAGGGGACCATCCTGTTGATGGAGGAAACATATGATAT 452  
 1760 CCATCATCTGGACTTACTTCACTTGGCCAGGTCACCTAGCGATGGTGGCACCTACGATAT 1819

453 CTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAGGGGATGGCCACATTTAAACAATA 512  
 1820 CTACTCAACCCAGCGTGTCAACACGCTTCCATTGAGGGAACCTTCCACCTTCAACAGTA 1879  
 513 TTGAGTGTTCGAAGATCGAAACGACAGTGGGACGATTTCTGTGACGACCACTTTAG 572  
 1880 CTGGTCACTTCCGACCGAGAACGAGTCCGGGGAACTGTCCACGCGGCAACCACTTTGC 1939  
 573 AGCTTGGGAAAACCTTAGGGATGAATATGGGAAAATGTATGAAGTCCGCGCTTACTGTAGA 632  
 1940 AGCATGGGAAGGCACTTGGACCTTGAATGGGCACTTATAACTATATGATTTGTGTCACCGA 1999  
 633 AGGCTATCAAAAGTAGCGGAAGTGTCTA 658  
 2000 AGGCTACGAGAGCAGTGGCTCTAGTA 2025

RESULT 14  
 US-10-213-990-64  
 ; Sequence 64, Application US/10213990  
 ; Publication No. US20030082595A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jiang, Bo  
 ; APPLICANT: Bussey, Howard  
 ; APPLICANT: Storms, Reg  
 ; APPLICANT: Roemer, Terry  
 ; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL  
 ; FILE REFERENCE: 10182-019-999  
 ; CURRENT APPLICATION NUMBER: US/10/213,990  
 ; CURRENT FILING DATE: 2002-08-05  
 ; NUMBER OF SEQ ID NOS: 72  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 64  
 ; LENGTH: 712  
 ; TYPE: DNA  
 ; ORGANISM: Aspergillus  
 ; US-10-213-990-64

Query Match 14.1%; Score 104.8; DB 14; Length 712;  
 Best Local Similarity 55.8%; Pred. No. 4.5e-20;  
 Matches 221; Conservative 0; Mismatches 172; Indels 3; Gaps 1;  
 270 TGGTAACATGTCCATAAACTACGGAGCCAACTTCCAAACCAATGGTAATGGCTATTATTG 329  
 304 TGGATTGAGCGGTCACTACAGCGGCTCTGGGAGACGCGGAAACGGCTACCTCTC 363  
 330 CGTCTATGTTGGACTGTTGACCCCTTCTGCGAATATTATATGTCGACAGTTGGGGCAA 389  
 364 CGGTACGGCTGGACGACCGTCCGCTGCTGCAATTTCTACATCGTGGAGAGTTACGGCTC 423  
 390 CTGGCGGTCCACCA---GGAGCAACGCTTAAGGGACCATCCTGTTGATGGGGAACATA 446  
 424 CTATGACCCCTCCACGGGAGCCACCATCTCGGCACCGCTGAGAGCGAGCGGGGCCACGTA 483  
 447 TGNATCTACGAGACTCTTTAGAGTCAATCAACCCCTCAATTAAGGGGATTTGCCACATTTAA 506  
 484 CAACCTCTACAGACGACCGGACGAGATGCGCGTCCATCCAGGCAACGGCTACTTTGA 543  
 507 ACAATATTGGAGTGTTCGAAGATCGAAACGACGAGTGGGACGATTTCTGTAGCAACCA 566  
 544 CCAGTACTGTCGGTTCGGACTTCGCAACCGGAGAGTGGAACTGTGACGACGAGAAACCA 603  
 567 CTTTAGAGCGTGGGAAAACCTTAGGGATGAATATGGGAAAATGTATGAAGTCCGCGCTAC 626  
 604 CTTTGTGCTGGAGAAATCGGGTCTGCAATTTGGGAACTTTGACTATATGATTGTTGC 663  
 627 TGTGAAGCTATCAAGTAGCGGAGTGTCTATGT 662  
 664 GACGGAGGGGTACGAGCAGCGGCTCTGCTACTAT 699

RESULT 15





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 10, 2004, 16:24:32 : Search time 3225.17 Seconds  
(without alignment)  
8406.125 Million cell updates/sec

Title: US-09-909-207-4  
Perfect score: 744  
Sequence: 1 ATGACACAAGAAATTCAC.....TAACCTTGGATAAAACAAT 744

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_g881:\*  
9: gb\_g882:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98.6	13.3	618	7	CF472462 RTDS1_9_C
2	93	12.5	728	7	CN132941 OX1_9_D10
3	93	12.5	772	7	CN133022 OX1_9_D10
4	91.8	12.3	744	7	CF867983 trico13xe
5	91.8	12.3	799	6	CB898036 trico13xe
6	86.6	11.6	738	6	CD464145 ETH1_48_B
7	83.4	11.2	921	6	CD458837 EG08_04B1
8	82.8	11.1	603	8	AQ160513 mgxb0006C
9	82.2	11.0	670	8	AQ447125 mgxb0005C
10	82.2	11.0	720	8	AQ361561 mgxb0004B
11	82.2	11.0	750	8	AQ160254 mgxb0003L
12	81.6	11.0	786	8	AQ325248 mgxb0021M
13	81.4	10.9	583	8	AQ399120 mgxb0001B
14	76.6	10.3	561	8	AQ396475 mgxb0010M
15	76.4	10.3	617	1	AJ638869 AJ638869
16	76.2	10.2	768	7	CF881056 trico083xj
17	76.2	10.2	822	6	CB907827 trico083xj
18	74.6	10.0	753	8	AQ448084 mgxb0016B
19	74.4	10.0	418	8	AQ398756 mgxb0005L
20	71.8	9.7	746	6	CB901964 trico28xi
21	71.8	9.7	746	7	CF871731 trico28xi
22	69.8	9.4	520	5	BQ165950 WHE0821-0
23	66.2	8.9	561	5	BQ471960 HV04A02r
24	63.2	8.5	473	6	CD464005 ETH1_48_B

C 25	60.4	8.1	483	5	BQ664593	HV04A02u
C 26	53.6	7.2	646	8	AQ449078	mgxb0022H
C 27	51.6	6.9	493	7	CO135442	EST830113
28	49.4	6.6	490	2	BZ200865	WHE0821-0
29	49.4	6.6	532	4	BM134812	WHE0452_F
30	48.4	6.5	532	7	CN008236	WHE0452_F
31	47.8	6.4	743	7	CF865758	WHE2638_H
32	47.8	6.4	802	6	CB895680	trico003XB
33	46.6	6.3	503	5	BQ166480	trico003XB
34	46.6	6.3	1582	9	CC820765	WHE0861_D
35	46.4	6.2	456	1	AJ637947	SSH_RfFD1
36	45.2	6.1	587	6	CD033274	AJ637947
37	44.8	6.0	614	6	CA278095	mgm012XD
38	41.6	5.6	488	4	BM135798	SCBFSDD203
39	41.6	5.6	591	4	BM135798	WHE2624_C
40	41.4	5.6	1201	9	CNS007JU	BJ375180
41	40.8	5.5	622	4	BJ401024	BJ375180
42	40.8	5.5	642	4	BJ401024	Drosophila
43	40.8	5.5	655	4	BJ372089	BJ401024
44	40.8	5.5	782	8	BH066221	BJ402691
45	40.6	5.5	474	4	BJ398920	BJ372089

ALIGNMENTS

RESULT 1  
CF472462  
LOCUS  
DEFINITION  
RTDS1\_9\_C03.g1\_A015 Drought-stressed loblolly pine roots DS1 Pinus  
taeda cDNA clone RTDS1\_9\_C03\_A015 5', mRNA sequence.  
ACCESSION  
CF472462  
VERSION  
CF472462.1  
KEYWORDS  
EST.  
SOURCE  
Pinus taeda (loblolly pine)  
ORGANISM  
Pinus taeda  
REFERENCE  
1 (bases 1 to 618)  
AUTHORS  
Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J.,  
Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and  
Neale, D.  
TITLE  
An EST database from drought-stressed loblolly pine (Pinus taeda)  
JOURNAL  
Unpublished (2003)  
COMMENT  
Other ESTs: RTDS1\_9\_C03\_b1\_A015  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@cuga.edu  
RNA prepared and library constructed by W. Walter Lorenz, School of  
Forestry, University of Georgia; plant material prepared at the  
University of Florida; sequencing done in the Laboratory for  
Genomics and Bioinformatics, University of Georgia. Sequence ends  
have been trimmed to exclude vector and regions below Phred quality  
16. Three-prime sequences are presented as their reverse complement  
and have been trimmed to exclude polyA.  
Seq primer: JFNREV (CAGGAACAGCTATGACC).

FEATURES  
source  
1..618  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="CCLONES"  
/db\_xref="taxon:3352"  
/clone="RTDS1\_9\_C03\_A015"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="Drought-stressed loblolly pine roots DS1"  
/notes="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The  
library was prepared from polyA+ RNA from drought-stressed  
loblolly pine (Pinus taeda) roots. Water was withheld from



COMMENT Other ESTs: OX1\_9\_D10.bi.A002  
 Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu  
 Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in  
 the Human Genome Center, University of Tokyo Institute of Medical  
 Science; plant material and RNA prepared at Texas A & M University;  
 sequencing done in the Laboratory for Genomics and Bioinformatics,  
 University of Georgia. Sequence ends have been trimmed to exclude  
 vector and regions below Phred quality 15. Three-prime sequences  
 are presented as their reverse complement and have been trimmed to  
 exclude polyA.  
 Seq primer: Sug5 (CTTCTGCTCTAAAAGTGGG).

## FEATURES

source  
 1..772  
 Location/Qualifiers  
 /organism="Sorghum bicolor"  
 /mol\_type="mRNA"  
 /cultivar="BTx623"  
 /db\_xref="taxon:4558"  
 /clone="OX1\_9\_D10.A002"  
 /lab\_host="DH10B-Ti phage-resistant E. coli"  
 /clone\_lib="Oxidatively-stressed leaves and roots"  
 /note="Organ: Leaf and Root; Vector: pME18S-FL3; Site 1:  
 XhoI; Site 2: XhoI; The library was prepared from polyA+  
 RNA from oxidatively stressed, hydroponically grown  
 sorghum seedlings. At 8 days of age, growth medium was  
 supplemented with hydrogen peroxide to 0.003% and leaves  
 were misted with 10 mM methyl viologen. Leaves and roots  
 were harvested at 3, 12 and 27 hr after treatment and all  
 tissue pooled. Double-stranded cDNA was cloned  
 unidirectionally into different DraIII sites of the  
 pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG,  
 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA  
 insert."

## ORIGIN

Query Match 12.5%; Score 93; DB 7; Length 772;  
 Best Local Similarity 53.9%; Pred. No. 3.9e-15;  
 Matches 214; Conservative 0; Mismatches 180; Indels 3; Gaps 1;  
 QY 281 CCATAAACTACGAGCAACTTCCAAACCAATGTAATGCGTATTTATGCGTATGTT 340  
 Db |||||  
 QY 324 CCATCAACTATGCGGTTCTTTCAGCCCTCAGGGTAACGGCTACCTCTGCGTCTACGGCT 383  
 Db |||||  
 QY 341 GGAATGTTGACCTCTTGTGCAATATATATTTGTCACAGTTGGGCA---ACTGGCGTC 397  
 Db |||||  
 QY 384 GGAATGCGAGCCCTCTCGTGTAGTACTACGTATCGAGAACTACGGCACTTACACCCCTG 443  
 Db |||||  
 QY 398 CACCAGGAGCAAGCGCTAAGGGGACCATCTCTGTATGAGGAACTATATATCTACG 457  
 Db |||||  
 QY 444 GCTCTGTGCGCAGCAGCAGGGCAGCGTCTAAGCAGCGGCACTACGATCTCTACC 503  
 Db |||||  
 QY 458 AGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTCGCAATTAACAATATGGA 517  
 Db |||||  
 QY 504 AGACCCCGCTACACACCGCCCTCTATCGAGCGCAACAGACCTTCAACAGTACTGGG 563  
 Db |||||  
 QY 518 GTGTTGGAAGATCGAAACGACAGTGGCAGGATTTCTGTACGACCAACCTTTAGAGCGT 577  
 Db |||||  
 QY 564 CCATCGCGCCCAACAGCGCAGCGCGCGCTCAACATGCAAGTATCTTCAATGCTT 623  
 Db |||||  
 QY 578 GGGAAACTTAGGGATGAATATGGGAAATGATGAGTGGCGCTTACTGTAGAAGGCT 637  
 Db |||||  
 QY 624 GGGCTAACGCTGTATGAGACTTGGAAAACCACTATACAGATCTCTGGCTACCGAGGAT 683  
 Db |||||  
 QY 638 ATCAAAGTACGGAAGTGTGTAATGTATATAGCAATAC 674  
 Db |||||  
 QY 684 ACCAGAGCAGTGGATCTTCTTCTATCTATGTGCCAGAC 720  
 Db |||||

## RESULT 4

CF867983  
 LOCUS CF867983  
 DEFINITION tricol3xe09.bl T.reesei mycelial culture, Version 6 October 2003  
 Hypocrea jecorina cDNA clone tricol3xe09, mRNA sequence.  
 ACCESSION CF867983  
 VERSION CF867983.1  
 KEYWORDS GI:38122635  
 SOURCE EST  
 ORGANISM Hypocrea jecorina (anamorph: Trichoderma reesei)  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.  
 REFERENCE 1 (bases 1 to 744)  
 AUTHORS Diener, S.E., Dunn-Coleman, N., Foreman, P., Houfek, T.D.,  
 Teunissen, P.J.M., van Solingen, P., Dankmeyer, L., Mitchell, T.K.,  
 Ward, M. and Dean, R.A.  
 TITLE Characterization of the protein processing and secretion pathways  
 in a comprehensive set of expressed sequence tags from Trichoderma  
 reesei  
 JOURNAL FEMS Microbiol. Lett. 230 (2), 275-282 (2004)  
 COMMENT Contact: Ralph A. Dean  
 Fungal Genomics Laboratory  
 North Carolina State University  
 Campus Box 7251, Raleigh, NC 27695, USA  
 Tel: 919-513-0020  
 Fax: 919-513-0024  
 Email: ralph.dean@ncsu.edu  
 Seq primer: LR-F1 primer.  
 FEATURES  
 source  
 1..744  
 Location/Qualifiers  
 /organism="Hypocrea jecorina"  
 /mol\_type="mRNA"  
 /strain="QM6a"  
 /db\_xref="taxon:51453"  
 /clone="tricol3xe09"  
 /dev\_stage="mycelia"  
 /clone\_lib="T.reesei mycelial culture, Version 6 October  
 2003"  
 /note="Vector: pREP3V, Site 1: Not 1/Sal I; Mycelial  
 culture grown from 24 hrs to 6 days with varying Carbon  
 and Nitrogen sources and concentrations."  
 ORIGIN  
 Query Match 12.3%; Score 91.8; DB 7; Length 744;  
 Best Local Similarity 55.0%; Pred. No. 8.5e-15;  
 Matches 202; Conservative 0; Mismatches 162; Indels 3; Gaps 1;  
 QY 274 AACATGTCATATAAATACGAGGCAACTTCCAAACCAATGTAATGCGTATTTATGCGTC 333  
 Db |||||  
 QY 353 AACAGGTCATCAACTTCTCGGCGAGCTACACCCCAACGGCAACAGCTACCTCTCCGTG 412  
 Db |||||  
 QY 334 TATGTTGAGCTGTGACCTCTTGTGCAATATATATTTGTCAGAGTTGGGGCACTGG 393  
 Db |||||  
 QY 413 TACGGCTGGTCCCGCAACCCCTGTATCGAGTACTACATCGTCGAGAACTTTGGCACCTAC 472  
 Db |||||  
 QY 394 CGTCCACCA---GGAGCAAGCGCTAAGGGACCATCTGTTGATGGAGCAATATGAT 450  
 Db |||||  
 QY 473 AACCCGTCACGGGCGCACCAAGTGGGCGAGGTCACTCCGACGGCAGCGCTTACGAC 532  
 Db |||||  
 QY 451 ATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTTGCCACATTTAAACAA 510  
 Db |||||  
 QY 533 ATTTACCGCAGCAGCGCGTCAACAGCGCTCCATCATCGGCACCGCCACCTTTTACCAG 592  
 Db |||||  
 QY 511 TATTGGAGTGTTCGAAGATCGAAACGACAGTGGCAGCATTTTGTTCAGCAACCACTTT 570  
 Db |||||  
 QY 593 TACTGTGCTCGCGCGCAACCCAGCTCGAGCGGCTCGTCAACACACGCGCAACCACTTC 652  
 Db |||||  
 QY 571 AGAGCGTGGAAAACTTAGGGATGATATATGGGAAATATGTAAGTCCGGCTTACTGTA 630  
 Db |||||  
 QY 653 AACCGGTGGGCTCAGCAAGGCTGAGCTCGGGACGATGATGATACAGATTGTTGCCGTG 712  
 Db |||||  
 QY 631 GAAGGCT 637  
 Db |||||

Db 713 GGAGGGT 719

RESULT 5  
CB898036  
LOCUS  
DEFINITION  
tricol3xe09 T. reesei mycelial culture, Version 3 april Hypocrea EST 02-JUL-2003  
ACCESSION  
CB898036  
VERSION  
tricol3xe09  
KEYWORDS  
J. Biol. Chem. 278 (34), 31988-31997 (2003)  
SOURCE  
Hypocrea jecorina (anamorph: Trichoderma reesei)  
ORGANISM  
Hypocrea jecorina  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Hypocrematidae; Hypocreales; Hypocreaceae; Hypocrea.  
REFERENCE  
1 (bases 1 to 799)  
Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Diener, S.,  
Dunn-Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J.,  
Kelley, A.S., Meerman, H.J., Mitchell, T., Mitchinson, C.,  
Olivares, H.A., Teunissen, P.J., Yao, J. and Ward, M.  
Transcriptional regulation of biomass-degrading enzymes in the  
filamentous fungus Trichoderma reesei  
J. Biol. Chem. 278 (34), 31988-31997 (2003)  
PUBMED  
12788920  
COMMENT  
Contact: Pamela K. Foreman  
Genencor Intl.  
925 Page Mill Road, Palo Alto, CA 94304, USA  
Tel: (650) 846-7635  
Fax: (650) 621-7817  
Email: pforeman@genencor.com  
Seq primer: LT-F1 primer.  
Location/Qualifiers  
1..799  
/organism="Hypocrea jecorina"  
/mol\_type="mRNA"  
/strain="QM6a"  
/db\_xref="taxon:51453"  
/clone="tricol3xe09"  
/dev\_stage="mycelia"  
/clone\_lib="T. reesei mycelial culture, Version 3 april"  
/note="Vector: pREP3Y; Site 1: Not I/Sal I; Mycelial  
culture grown from 24 hrs to 6 days with varying Carbon  
and Nitrogen sources and concentrations."

FEATURES  
source  
1..799  
Location/Qualifiers  
1..799  
/organism="Sorghum bicolor"  
/mol\_type="mRNA"  
/cultivar="BTx623"  
/db\_xref="taxon:4558"  
/clone="ETH1\_48\_B06\_A002"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="Ethylene-treated seedlings"  
/note="Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The  
library was prepared from polyA+ RNA from seedlings grown  
in hydroponic culture. At 8 days of age, medium was  
supplemented with 0.5 mM 1-aminocyclopropane-1-carboxylic  
acid (ACC) to induce endogenous ethylene (ETH) production.  
Roots and shoots were harvested after 27 and 72 hr and  
material from both time points was combined prior to RNA  
isolation. Double-stranded cDNA was cloned  
unidirectionally into different DralII sites is CACTGTGTC,  
pME18S-FL3 vector (5-prime DralII site is CACTGTGTC,  
3-prime DralII site is CACCATGTG). XhoI excises the cDNA  
insert."

ORIGIN  
Query Match 11.6%; Score 86.6; DB 6; Length 738;  
Best Local Similarity 53.5%; Pred. No. 2.6e-13;  
Matches 204; Conservative 0; Mismatches 174; Indels 3; Gaps 1;  
QY 281 CCATAAAGTACGGAGCCAACTTCCAAATGTAATGCGTATTATGCTCTATGTT 340  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 631 GAAGGCT 637  
Db 768 GGAGGGT 774

RESULT 6  
CD464145  
LOCUS  
DEFINITION  
ETH1\_48\_B06\_g1\_A002 Ethylene-treated seedlings Sorghum bicolor cDNA  
ACCESSION  
CD464145  
VERSION  
CD464145.1  
KEYWORDS  
EST.  
SOURCE  
Sorghum bicolor (sorghum)  
ORGANISM  
Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Sorghum.  
REFERENCE  
1 (bases 1 to 738)  
Cordonnier-Pratt, M.-M., Wentzel, V., Suzuki, Y., Sugano, S.,  
Klein, R.R., Liang, C., Sun, F., Sullivan, R., Shah, M., Salzman, R.,  
Chua Tan, N., Gonzalez, M., Lane, S., Miller, V., Nanda, P.,  
Olaseinde, O., Eastman, A. and Pratt, L.H.  
An EST database from Sorghum: 1-aminocyclopropane-1-carboxylic acid  
(ACC)-treated seedlings  
Unpublished (2003)  
JOURNAL  
Other ESTs: ETH1\_48\_B06.bl A002  
COMMENT  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
Library constructed by Dr. Yutaka Suzuki and Dr. Sunio Sugano in  
the Human Genome Center, University of Tokyo Institute of Medical  
Science; plant material and RNA prepared at Texas A & M University;  
sequencing done in the Laboratory for Genomics and Bioinformatics,  
University of Georgia. Sequence ends have been trimmed to exclude  
vector and regions below phred quality 16. Three-prime sequences  
are presented as their reverse complement and have been trimmed to  
exclude polyA.  
Seq primer: Sug5 (CTTCTGCTCTAAAGCTGCG).  
Location/Qualifiers  
1..738  
/organism="Sorghum bicolor"  
/mol\_type="mRNA"  
/cultivar="BTx623"  
/db\_xref="taxon:4558"  
/clone="ETH1\_48\_B06\_A002"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="Ethylene-treated seedlings"  
/note="Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The  
library was prepared from polyA+ RNA from seedlings grown  
in hydroponic culture. At 8 days of age, medium was  
supplemented with 0.5 mM 1-aminocyclopropane-1-carboxylic  
acid (ACC) to induce endogenous ethylene (ETH) production.  
Roots and shoots were harvested after 27 and 72 hr and  
material from both time points was combined prior to RNA  
isolation. Double-stranded cDNA was cloned  
unidirectionally into different DralII sites is CACTGTGTC,  
pME18S-FL3 vector (5-prime DralII site is CACTGTGTC,  
3-prime DralII site is CACCATGTG). XhoI excises the cDNA  
insert."

ORIGIN  
Query Match 11.6%; Score 86.6; DB 6; Length 738;  
Best Local Similarity 53.5%; Pred. No. 2.6e-13;  
Matches 204; Conservative 0; Mismatches 174; Indels 3; Gaps 1;  
QY 281 CCATAAAGTACGGAGCCAACTTCCAAATGTAATGCGTATTATGCTCTATGTT 340  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 345 CCATCAAGTACTCTGTTTACAACTCAACGGAACAGCTACCTCGCTGTTTACGGAT 404  
 QY 341 GGACTGTTGACCTCTGTTGCAATATTATATGTCGACAGTTGGGGCACTGGCGTCCAC 400  
 Db 405 GGACTCAGAACCTCTCATCGAGTACTACATCGTTGAGAACTTCGGCACTTACAACTCCCT 464  
 QY 401 CA---GGAGCAACGCTTAAGGGACCACTGTTGATGGGGAACATATGATATCTACG 457  
 Db 465 CTTCGGGGCGGCAGAAAGGCTGAGGTCACCTGTTGACGGATCTGTTGACGATCTACG 524  
 QY 458 AGACTCTTAGAGTCAATCAACCTCCATTAAAGGGATTGCCACATTTAAACAATATGGA 517  
 Db 525 TCAGCAACCGTGTCAACGCCCTCCATTGAGGGTAACAAGACCTTTTCAGCAGTCTCGT 584  
 QY 518 GTGTTGCAAGATGCAAAACGACGAGTGGCAAGATTTCTGTGACGAACCACTTTAGAGCGT 577  
 Db 585 CTGTTGCAAGCAACAAAGCGATCCAGCGATCCGTCGATCAACACCGGTCTCCTTCCAGGCT 644  
 QY 578 GGGAAACTTAGGGATGAATATGGGGAATAATGATGAAGTCGCGCTTACTGTAGAGGCT 637  
 Db 645 GGAATAATGTCGGCTCAACCTTTGGTAACCAACTATCAGATCCTTGTGTTGAGGCT 704  
 QY 638 ATCAAAGTAGCGAAGTGCTA 658  
 Db 705 ACTACAGCTCTGGCTCGCA 725

RESULT 7  
 CD458837 921 bp mRNA linear EST 14-JUN-2004  
 LOCUS Fg08\_04b10\_A Fg08\_AAPC\_ECORC\_Fusarium\_graminearum\_complex\_substrate  
 DEFINITION Gibberella zeae cDNA clone Fg08\_04b10, mRNA sequence.  
 CD458837  
 CD458837.1 GI:31373577

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Gibberella zeae  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.  
 1 (bases 1 to 921)  
 Hattori, J., Heys, R., Chapados, J., Couroux, P., Harris, L.J.,  
 Watson, R.J., Lacroix, C., Ouellet, T., Robert, L.S., Singh, J.A.,  
 Spott, B., and Tinker, N.A.  
 A cDNA library prepared from Fusarium graminearum grown on a  
 complex plant substrate  
 Unpublished (2003)  
 Contact: Watson, Robert J.  
 Eastern Cereal and Oilseed Research Centre  
 Agriculture and Agri-food Canada  
 Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,  
 CANADA  
 Tel: (613) 759-1655  
 Fax: (613) 759-1701  
 Email: watsonrj@agr.gc.ca

FEATURES  
 source  
 1..921  
 Location/Qualifiers  
 /organism="Gibberella zeae"  
 /mol\_type="mRNA"  
 /strain="DAOM 180378"  
 /db\_xref="taxon:5518"  
 /clone="Fg08\_04b10"  
 /tissue\_type="Mycelium"  
 /dev\_stage="Asexual"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="Fg08\_AAPC\_ECORC\_Fusarium\_graminearum\_complex\_s  
 ubstrate"  
 /notes="Vector: pBluescript II+; Site\_1: EcoRI; Site\_2:  
 XhoI; Fusarium graminearum grown on a complex plant  
 substrate-- wheat leaves treated to remove most of the low  
 molecular weight, water-soluble components."

Query Match 11.2%; Score 83.4; DB 6; Length 921;

Best Local Similarity 52.4%; Pred. No. 2.2e-12;  
 Matches 208; Conservative 0; Mismatches 186; Indels 3; Gaps 1;

QY 281 CCATAAAGTACGGAGCAACTTCCCAACCAATGTAATCGTATTTATGCGTCTATGTT 340  
 Db 363 CCATCAACTACGGAGGTTCTTTCAACCTCAGGGTAACGGATACCTTTGCGTTTACGGAT 422  
 QY 341 GGACTGTTGACCTCTGTTGCAATATTATATGTCGACAGTTGGGGCACTGGCGTCCAC 400  
 Db 423 GGACCCGGGTCCCTCGTTCGAGTACTACGTCATCGAGAGTTACGGTTCTTTACAACCCG 482  
 QY 401 CAGGA---GCAACGCCCTTAAGGGACCACTCATCTGTTGATGGGGAACATATGATATCTACG 457  
 Db 483 GCAGCCAGGCTCAGCAGCGAGGTACCGTCTACACCGACGGTGACACCTTACGATCTCTATA 542  
 QY 458 AGACTCTTAGAGTCAATCAACCTCCATTAAAGGGATTGCCACATTTAAACAATATGGA 517  
 Db 543 TGTCCACCCGTTACACAGCGCTTCGATCGACGGTTCAGACCTTCAACCACTTACGATCTGT 602  
 QY 518 GTGTTGCAAGATGCAAAACGACGAGTGGCAAGATTTCTGTGACGAACCACTTTAGAGCGT 577  
 Db 603 CCATCCGCGCAACAGCGGTACCGTCCGTCACATCGAGAACCACCTTCAATGCTT 662  
 QY 578 GGGAAACTTAGGGATGAATATGGGGAATAATGTAAGTCCGCTTACTGTAGAGGCT 637  
 Db 663 GGAGATCTGCTGGCATGAACCTCGGAAACCACTTACTACGATTTCTGCCACTGAGGGTT 722  
 QY 638 ATCAAAGTAGCGAAGTGCTAATGTTATATAGCAATAC 674  
 Db 723 ACCAGAGCAGTGGCTCATCTTCTATCTATGTCAGAC 759

RESULT 8  
 AQ160513/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Magnaporthe grisea (anamorph: Pyricularia grisea)  
 Magnaporthe grisea  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.  
 1 (bases 1 to 603)  
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,  
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.  
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea  
 Genome  
 Unpublished (1998)  
 Contact: Dean RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson University, Clemson, SC 29634  
 Tel: 864 656 5737  
 Fax: 864 656 4293  
 Email: rdean@clemson.edu  
 Seq primer: GGAACAGCTATGACCATG  
 Class: BAC ends  
 High quality sequence stop: 424.  
 Location/Qualifiers

FEATURES  
 source  
 1..603  
 /organism="Magnaporthe grisea"  
 /mol\_type="genomic DNA"  
 /strain="70-15"  
 /db\_xref="taxon:148305"  
 /clone="mgxb0006C21r"  
 /tissue\_type="Protoplasts"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="CUGI Rice Blast BAC Library"  
 /notes="Vector: pBACWICH; Site\_1: HindIII; Site\_2: HindIII;  
 Rice blast is one of the most devastating fungal diseases  
 of rice world wide. It is a filamentous ascomycete with



**TITLE** A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome  
**JOURNAL** Unpublished (1998)  
**COMMENT** Contact: Dean RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson University, Clemson, SC 29634  
Tel: 864 656 5737  
Fax: 864 656 4293  
Email: rdean@clemson.edu  
Seq primer: GGAAACAGCTATGACCATG  
Class: BAC ends  
High quality sequence stop: 321.

**FEATURES**  
**SOURCE**

```

1. 1720
/organism="Magnaporthe grisea"
/mol_type="genomic DNA"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0004B19r"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice Blast BAC Library"
/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

```

## ORIGIN

Query Match	11.0%;	Score 82.2;	DB 8;	Length 720;
Best Local Similarity	54.1;	Pred. No. 4.5e-12;		
Matches 236;	Conservative 0;	Mismatches 189;	Indels 11;	Gaps 3;
QY	229	AAAGGTAAAAATTC	CAATGAAACACAACACACACCAACCAAGTTGGTAACTATGTCCATAAAC	288
DB	445	AAAGGAAAAAAGAGACT	AAACAACAACAACAAAAA-----CAGCGCGTCTATCAAC	391
QY	289	TACGGAGCCAACTT	CCAAACAAATGGTAAATGCGTATTATTCGCTCTATGTTGGACTGTT	348
DB	390	TACTCGGGCAGCTA	CAGCCNCAGGGCAACTCATACCTGGCCGCTACGGCTGACGGCGC	331
QY	349	GACCCCTCTTGT	CGAATATTAATTGTGACAGATTTGGGGCAACTGGCGTCCACCA---GGA	405
DB	330	AACCCGCTGAT	CGAGTACTACGTGTGGAGAGCTTTGGCAGCTACAACCGTCTGTCGGGC	271
QY	406	GCAACGCCCTA	AGGGGACCATCACTGTTTGATGGAGAGAACATATGATATCTACGAGACTCTT	465
DB	270	GCCACCAACCG	CGGCTCTTCACTCTCGGACGCGCAGCACCTACGACATCTCTGGTCAGCACCC	211
QY	466	AGAGTCAATCAA	CCCTCCATTTAAGGGGATTCGCCACATTTTAAACAATATTGGAGTGTTCGA	525
DB	210	CGTACAAACG	ACCCCTCCATCGACGGCACCAAGACCTTTTACGAGTCTCTGGTCGGTGGCG	151
QY	526	AGATCGAAA	CGCAGAGTGGCAGCATTTCTGTGACGCAACCACTTTTAGAGCGTGGGAAAAC	585
DB	150	CGAAACAGCG	CGCCAGGGCACCGTCACTTTTGCCAAACACAGCTCAAGCCTTGGGGCAAC	91
QY	586	TTAGGGATGA	TATGGGAAAAATGTATGAAGTTCGCG---CTTACTGTAGAAAGGTATCAA	642
DB	90	GCCGSCCTCA	ACCTCGGCAACCAAGTGGAACTACACAGATCCTGGCCGTCAGGGGTACCAC	31
QY	643	AGTAGCGGA	AGTGTCTA 658	
DB	30	AGCAGCGCT	CCGCCA 15	

RESULT 11	ACCESSION
AQ160254/c	VERSION
LOCUS	KEYWORDS
DEFINITION	SOURCE
	ORGANISM

REFERENCE  
AUTHORS

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1965, Vol. 68, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Technology	1978, Vol. 1, No. 2, pp. 1-10
3. The Importance of Parental Involvement	Journal of Educational Research	1985, Vol. 88, No. 3, pp. 1-10
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1990, Vol. 93, No. 4, pp. 1-10
5. The Role of the School in the Community	Journal of Educational Research	1995, Vol. 98, No. 5, pp. 1-10
6. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	2000, Vol. 103, No. 6, pp. 1-10
7. The Role of the Teacher in the Classroom	Journal of Educational Research	2005, Vol. 108, No. 7, pp. 1-10
8. The Impact of Technology on Education	Journal of Educational Technology	2010, Vol. 3, No. 8, pp. 1-10
9. The Importance of Parental Involvement	Journal of Educational Research	2015, Vol. 118, No. 9, pp. 1-10
10. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	2020, Vol. 123, No. 10, pp. 1-10

## FEATURES

**FEATURES**

## ORIGIN

Query Match	Best Local Match
QY	
Db	
QY	
Db	
QY	
Db	
QY	
Db	
QY	

AQ160254 750 bp DNA linear GSS 09-SEP-1998  
mgxb0003L19r CUGI Rice Blast BAC Library Magnaporthe grisea genomic  
clone mgxb0003L19r, genomic survey sequence.  
AQ160254  
AQ160254.1 GI:3557243  
GSS.  
Magnaporthe grisea (anamorph: Pyricularia grisea)  
Magnaporthe grisea  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.  
1 (bases 1 to 750)  
Yu, X., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,  
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.

Genome  
Unpublished (1998)  
Contact: Dean RA  
Clemson University  
Clemson University  
Clemson University  
100 Jordan Hall, Clemson University, Clemson, SC 29634  
Tel: 864 656 5737  
Fax: 864 656 4293  
Email: rdean@clemson.edu  
Seq primer: GGAAACAGCTATGACCATG  
Class: BAC ends  
High quality sequence stop: 344.

```

1. .750
/organism="Magnaporthe grisea"
/mol_type="genomic DNA"
/scrain="70-15"
/db_xref="taxon:148305"
/clone="mgx00003L19r"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice Blast BAC Library"
/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."

```

ch	11.0%; Score 82.2; DB 8; Length 750;
	1 Similarity 54.1%; Pred. No. 4.5e-12; Indels 11; Gaps 3;
	236; Conservative 0; Mismatches 189;
29	AAAGGTAATAAATTCATGAACACAAAACACCAACAAGTTGGTAACTGTCCATAAAC 288
45	AAAGGAAAAAAAAGAGACTAAACAACAACAACAAAA-----CAGCGCGGTTCATCAAC 391
89	TAGCGAGCCNACTTCCAAACCAATGGTAAATCGGTATTATTCGGTCTATGTTTGACTGTT 348
90	TACTCGGGGAGCTTACAGCCNCAGGGCNACTCATCTCGGCCGTCTACGGCTGACGCGGC 331
49	GACCCTCTTGTGCAATATTATATTGTCGACAGTTGGGGCAAATCTGGCGTCCACCA---GGA 405
30	AACCCGCTGATCAGTACTACGTGGTGGAGAGCTTTGGCAGCTACAACCCGTCGTCGGGC 271
06	GCAACCCCTAAGGGGACCATCACTGTTGTATGGAGGAACATATGATATCTACGAGACTCTT 465
70	GCCACCAACCGCGGTCTTTCACCTCGGACGGCAGCACCTTACGACATCTCTGTCAGCACC 211
66	AGAGTCAATCAACCCCTCCATTAAAGGGGATTGCCACATTTTAAACAANTATTGGAGTTTCGA 525



Db 210 CGCTAACACAGCCCTCCATCGACGGCACCAAGACCTTTACAGATTCTGGTGGTGGC 151

Qy 526 AGATCGAAGCCAGCAGTGGCAGGTTCTGTGACACACCTTTAGAGCGTGGAAAC 585

Db 150 CGCAACAGCGCGGCGGCGGACCGTCACTTTGCCAACACCGTCAACGCGCTGGCGCAAC 91

Qy 586 TTAGGGATGAATATGGGAAAAATGTATGAAGTCCGG---CTTACTGTAGAAGGCTATCAA 642

Db 90 GCGGCTCACTCGGCAACGATGGGAATACAGATCCTGGCGTGGAGGCTACCAAC 31

Qy 643 AGTAGCGGAAGTGCTA 658

Db 30 AGCAGCGGCTCGGCCA 15

RESULT 12

LOCUS AQ325248/c

DEFINITION mgxb0021M08r CUGI Rice Blast DNA linear GSS 08-JAN-1999

ACCESSION clone mgxb0021M08r, genomic survey sequence.

VERSION AQ325248

KEYWORDS GSS.

SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)

ORGANISM Magnaporthe grisea

REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.

AUTHORS 1 (bases 1 to 786)

Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.

TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome

JOURNAL Unpublished (1998)

COMMENT Contact: Dean RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson University, Clemson, SC 29634

Tel: 864 656 5737

Fax: 864 656 4293

Email: rdean@clemson.edu

Seq primer: GGAAACAGCTATGACCATG

Class: BAC ends

High quality sequence start: 49

High quality sequence stop: 468.

Location/Qualifiers

FEATURES

source

1..786

/organism="Magnaporthe grisea"

/mol\_type="genomic DNA"

/strain="70-15"

/db\_xref="taxon:148305"

/clone="mgxb0021M08r"

/tissue\_type="Protoplasts"

/lab\_host="E. coli DH10B"

/clone\_lib="CUGI Rice Blast BAC Library"

/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

ORIGIN

Query Match 11.0%; Score 81.6; DB 8; Length 786;

Best Local Similarity 54.1%; Pred. No. 6.8e-12;

Matches 236; Conservative 0; Mismatches 189; Indels 11; Gaps 3;

Qy 229 AAAGTAAATAATTCATGAAACACAAACACCAACCAAGTGGTGAACATGTCCATAAAC 288

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 500 AAAGGAAAAAAGAGACTAACACAAACAAACAAAAA-----CAGCCGCGTCAATCAAC 446

Qy 289 TAGGAGCCCACTTCCAAACCAAAATGTAATGCGTATTTATGCGTCTATGTTGACTGTT 348

Db 445 TACTCGGCGAGCTACAGCCCCCAGGGCAACTCATACCTGGCCGCTTACGGCTGGACGGC 386

Qy 349 GACCTCTTGTTCGAATATATTTGTGCACAGTTTGGGGCAACTGGCGTCCACCA---GGA 405

Db 385 AACCCGCTGATCGAGTACTACGTGGTGAGAGCTTTTGGCAGCTACAACCCGCTCGTGGGC 326

Qy 406 GCAACGCTAAGGGGACCATCATCTGTTGTATGGAGGAACATATGATATCTAGGACTCTT 465

Db 325 GCCACCAACGCGGGTCTTTCACTCGGACGCGAGCACCTACGACATCTCTGGTCAGACC 266

Qy 466 AGAGTCAATCAACCTCCATTAAGGGGATTGCCACATTTAAACAATATGGAGTGTTCGA 525

Db 265 CGTTACAACACAGCCCTCCATCGACGGCACCAAGACCTTTGAGCGTCTGGTGGTGGC 206

Qy 526 AGATCGAAGCCAGCAGTGGCAGGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAAC 585

Db 205 CGCAACAAGCGCGCCAGCGGCACCGTCACTTTTGGCAACCAACGCTCAACGCGCTGGCGCAAC 146

Qy 586 TTAGGGATGAATATGGGAAAAATGTATGAAGTCCGG---CTTACTGTAGAAGGCTATCAA 642

Db 145 GCGGCGCTCAACCTCGGCAACCAAGTGGAACTACAGATCTCTGGCCGCTCGAGGGCTACCAC 86

Qy 643 AGTAGCGGAAGTGCTA 658

Db 85 AGCAGCGGCTCGGCCA 70

RESULT 13

LOCUS AQ399120/c

DEFINITION mgxb0001B24f CUGI Rice Blast BAC Library Magnaporthe grisea genomic

clone mgxb0001B24f, genomic survey sequence.

ACCESSION AQ399120

VERSION AQ399120.1

KEYWORDS GI:4370147

SOURCE GSS.

ORGANISM Magnaporthe grisea (anamorph: Pyricularia grisea)

Magnaporthe grisea

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.

REFERENCE 1 (bases 1 to 583)

Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,

Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.

A BAC End Sequencing Framework to Sequence the Magnaporthe grisea

Genome

Unpublished (1998)

Contact: Dean RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson University, Clemson, SC 29634

Tel: 864 656 5737

Fax: 864 656 4293

Email: rdean@clemson.edu

Seq primer: TAATACGACTCACTATAGGG

Class: BAC ends

High quality sequence stop: 411.

Location/Qualifiers

1..583

/organism="Magnaporthe grisea"

/mol\_type="genomic DNA"

/strain="70-15"

/db\_xref="taxon:148305"

/clone="mgxb0001B24f"

/tissue\_type="Protoplasts"

/lab\_host="E. coli DH10B"

/clone\_lib="CUGI Rice Blast BAC Library"

/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice





COMMENT

Contact: Keon J  
Plant Pathogen Interactions Division,  
Rothamsted Research,  
Harpenden, Herts, UNITED KINGDOM  
Tel: +44(0)1582 763133  
Fax: +44(0)1582 760981  
Email: john.keon@bbsrc.ac.uk  
Insert Length: 800 Std Error: 100.00  
Seq primer: M13 reverse.  
Location/Qualifiers  
1..617  
/organism="Mycosphaerella graminicola"  
/mol\_type="mRNA"  
/strain="Strit"  
/db\_xref="taxon:54734"  
/clone="mgcl2d03f"  
/clone\_lib="Mgc"  
/note="Vector: pSPORT1; Library constructed from senescent  
wheat leaves 21-25 days after infection with  
Mycosphaerella graminicola exhibiting abundant hyphal  
growth and asexual sporulation"

FEATURES

source

ORIGIN

Query Match	10.3%	Score 76.4;	DB 1;	Length 617;
Best Local Similarity	55.6%	Pred. No. 1.9e-10;		
Matches 168;	Conservative	0;	Mismatches 131;	Indels 3; Gaps 1;

  

QY	281	CCATAACTACGAGCCAACTCCAAACCAATGGTAATGCGTATTTATGCGTCTATGTTT	340
Db	305	CCATCAACTACTCCGGCTCTTCAACCCCTCGGCAACGGCTACATCGCGCTACGGCT	364
QY	341	GGACTGTTGACCTCTTGTGNAATATATATGTCACAGTTGGGCAACTGGCGTCCAC	400
Db	365	GGACCCGCAACCCATGTGATCGAATACTACGTGTCGAATCTTACGGCACCTTCAACCCCA	424
QY	401	---CAGGAGCAACGCTTAAGGGGACCATCACTGTTGATGAGGAAACATATGATATCTACG	457
Db	425	GCTCTGCGGCTCAGAAGAGGGAACCGTCACACCGCGCGGACCTACGATATCTCTCC	484
QY	458	AGACTCTTAGAGTCAATCAACCTCCATTAGGGGATTGCCACATTTAAACATATTGGA	517
Db	485	AAACCAACCGGTGTCAACCAAGCCCTCCATTGACGCAACGAGACCTTTTCAGCAACTTCTGGA	544
QY	518	GTGTTGGAAGATCGAAACGACAGTGGCAGATTTCTGTGAGCAACCACTTTAGAGCGT	577
Db	545	CGTGGCGGACGAGAAGCGCGTCCGGCGGACGGTACCATGAAGAACCACTTCGACGCTC	604
QY	578	GG	579
Db	605	GG	606

Search completed: November 11, 2004, 01:16:49  
Job time : 3227.17 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 15:57:32 ; Search time 7319.01 Seconds  
(without alignments)  
9775.823 Million cell updates/sec

Title: US-09-909-207-10  
Perfect score: 1513  
Sequence: 1 AAATTGAATTGGTATATCT.....TGAACACCTCGTCACTAG 1513

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1513	100.0	1513	6	A48231 Sequence 10
2	1513	100.0	1513	6	A48232 Sequence 11
3	1513	100.0	1513	6	AR193055 Sequence
4	1513	100.0	1513	6	AR193056 Sequence
5	744	49.2	744	6	A48225 Sequence 4
6	744	49.2	744	6	A48226 Sequence 5
7	744	49.2	744	6	AR193051 Sequence
8	744	49.2	744	6	AR193052 Sequence 1
9	713.6	47.2	744	6	A68006 Sequence 1
10	713.6	47.2	744	6	AR163110 Sequence
11	673.6	44.5	744	6	A45313 Sequence 18
12	673.6	44.5	744	6	AR117325 Sequence
13	663	43.8	663	6	A48222 Sequence 1
14	663	43.8	663	6	A48223 Sequence 2
15	663	43.8	663	6	AR193049 Sequence
16	663	43.8	663	6	AR193050 Sequence
17	619	40.9	619	6	A48233 Sequence 12
18	619	40.9	619	6	AR193057 Sequence
19	587.2	38.8	744	6	A68016 Sequence 11

20	587.2	38.8	744	6	AR163117	Sequence
21	369.6	24.4	2246	1	AB029319	Bacillus
22	368	24.3	1531	1	BSXNY	X59059 Bacillus sp
23	288.2	19.0	1070	1	BPXYNA	X00660 Bacillus pu
24	282.8	18.7	1022	6	A42251	Sequence 1
25	282.8	18.7	1022	6	A42285	Sequence 35
26	282.8	18.7	1022	6	AR127019	Sequence
27	282.8	18.7	1022	6	AR127049	Sequence
28	282.8	18.7	1022	6	AR220023	Sequence
29	282.8	18.7	1022	6	AR220053	Sequence
30	282.8	18.7	1022	6	AR221306	Sequence
31	282.8	18.7	1022	6	AR221336	Sequence
32	281.8	18.6	1454	1	CLOXNB	M31726 Clostridium
33	281	18.6	5774	1	AB010958	AB010958 Clostridi
34	280.2	18.5	1011	1	AF490981	Bacillus
35	279	18.4	4128	1	AF047761	AF047761 Clostridi
36	277.6	18.3	687	1	AY526092	AY526092 Bacillus
37	275.4	18.2	1789	1	AF326785	AF326785 Bacillus
38	275	18.2	3493	1	CST508403	AJ508403 Clostridi
39	274.4	18.1	684	1	AF220528	AF220528 Bacillus
40	272.8	18.0	681	6	A42280	A42280 Sequence 30
41	272.8	18.0	681	6	A42282	A42282 Sequence 32
42	272.8	18.0	681	6	AR127045	AR127045 Sequence
43	272.8	18.0	681	6	AR127046	AR127046 Sequence
44	272.8	18.0	681	6	AR220049	AR220049 Sequence
45	272.8	18.0	681	6	AR220050	AR220050 Sequence

ALIGNMENTS

RESULT 1  
A48231 LOCUS A48231 1513 bp DNA linear PAT 07-MAR-1997  
DEFINITION Sequence 10 from Patent EP0698667.  
ACCESSION A48231  
VERSION A48231.1 GI:2302078  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 1513)  
AUTHORS De,B.E., Lahaye,A., Ledoux,P. and Detroz,R.  
TITLE Xylanase, microorganisms for its production, DNA molecules, process of preparation and use thereof  
JOURNAL Patent: EP 0698667-A 10 28-FEB-1996;  
COMMENT SOLVAY (BE)  
Other publication BE 1008751 960702  
Other publication BE 1008570 960604  
Other publication BR 9503454 960305  
Other publication JP 8092284 960409  
Other publication FI 953578 960127  
Other publication CA 2154628 960127  
Other publication AU 2508695 960208.

FEATURES

source  
1..1513  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 1513; DB 6; Length 1513;  
Best Local Similarity 100.0%; Pred. No. 1.5e-295;  
Matches 1513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AAATTCGAATGTGTATCTAATGATGATGACAAATCGTCACGTGTTTAACTAATCTC 60  
Db 1 AAATTCGAATGTGTATCTAATGATGATGACAAATCGTCACGTGTTTAACTAATCTC 60  
Qy 61 AAACCAATCTCTTTTATTTAAGCGTCAACCTTGAACCTTCAATCTTATCAAGAACATCTTT 120  
Db 61 AAACCAATCTCTCTTTTATTTAAGCGTCAACCTTGAACCTTCAATCTTATCAAGAACATCTTT 120

Qy	121	ATAGGA	ACTTTTCC	ATTTCG	CAAGAC	GATAAAAA	AAACCTTTTCC	CTTAATTTTATCT	TATCTG	180		
Db	121	ATAGGA	ACTTTTCC	ATTTCG	CAAGAC	GATAAAAA	AAACCTTTTCC	CTTAATTTTATCT	TATCTG	180		
Qy	181	CCTTGAT	CGGTTAA	TTTGTAA	AACTTTAT	TTTATAGTTT	ACGTGAT	GTTCCTCAT	TATCATAC	240		
Db	181	CCTTGAT	CGGTTAA	TTTGTAA	AACTTTAT	TTTATAGTTT	ACGTGAT	GTTCCTCAT	TATCATAC	240		
Qy	241	CATTAA	TATCAG	TTTAA	CGCTAG	AGTCA	TCTTTTTC	GGTTCTCA	AAAAATAC	300		
Db	241	CATTAA	TATCAG	TTTAA	CGCTAG	AGTCA	TCTTTTTC	GGTTCTCA	AAAAATAC	300		
Qy	301	ATTTAT	GTGCAT	ATTTTCT	CA	CGCGCT	CCATA	TGGAATAT	ATATATCT	360		
Db	301	ATTTAT	GTGCAT	ATTTTCT	CA	CGCGCT	CCATA	TGGAATAT	ATATATCT	360		
Qy	361	TTAAGT	AAATTTAG	TATATCT	TGCGTT	TATCA	AAATGTG	AGATAATCT	TAATTTGAT	420		
Db	361	TTAAGT	AAATTTAG	TATATCT	TGCGTT	TATCA	AAATGTG	AGATAATCT	TAATTTGAT	420		
Qy	421	AGCAG	CTATCC	AAAAAC	ACTG	ATGTTG	ACTCTT	TAAAG	AAAGTGTCT	ATCTAT	480	
Db	421	AGCAG	CTATCC	AAAAAC	ACTG	ATGTTG	ACTCTT	TAAAG	AAAGTGTCT	ATCTAT	480	
Qy	481	GATAA	TTATCC	AGTTT	CA	AAATTTG	AAATAG	TGTATG	GAATAGTT	TGAATG	540	
Db	481	GATAA	TTATCC	AGTTT	CA	AAATTTG	AAATAG	TGTATG	GAATAGTT	TGAATG	540	
Qy	541	CTGTG	AAAGG	AGGTAG	GTAC	CGTAG	ACTTC	ATTAC	CA	AAAAATTTAG	TTGTG	600
Db	541	CTGTG	AAAGG	AGGTAG	GTAC	CGTAG	ACTTC	ATTAC	CA	AAAAATTTAG	TTGTG	600
Qy	601	TTAAA	GAGGA	ATGCC	TATG	ACAC	AAAAAG	AAATTC	GACGTT	GAATTTAG	CTTTT	660
Db	601	TTAAA	GAGGA	ATGCC	TATG	ACAC	AAAAAG	AAATTC	GACGTT	GAATTTAG	CTTTT	660
Qy	661	TTGTTT	TGCA	TAACTTT	AC	CTG	CAG	AAATTA	ATTCAG	GCA	CAATTC	720
Db	661	TTGTTT	TGCA	TAACTTT	AC	CTG	CAG	AAATTA	ATTCAG	GCA	CAATTC	720
Qy	721	CATTG	GC	CAAC	CGAT	GGCTAT	GAATTTT	TG	AAAGATAG	CGGTG	CTGG	780
Db	721	CATTG	GC	CAAC	CGAT	GGCTAT	GAATTTT	TG	AAAGATAG	CGGTG	CTGG	780
Qy	781	AATG	ATCTCA	ATCA	TG	CGGTG	CTGTT	CAGTGC	CAATG	GAA	CAATG	840
Db	781	AATG	ATCTCA	ATCA	TG	CGGTG	CTGTT	CAGTGC	CAATG	GAA	CAATG	840
Qy	841	ATTCG	TAAAG	GTAAAA	ATTTCA	ATGAA	CA	CAAC	ACAC	CAAC	CAAGTGT	900
Db	841	ATTCG	TAAAG	GTAAAA	ATTTCA	ATGAA	CA	CAAC	ACAC	CAAC	CAAGTGT	900
Qy	901	CATAA	ACTAC	GGAG	CCAACTTT	CCAA	CCAAAT	GGTAAT	TTTAT	TGCGGT	CTATG	960
Db	901	CATAA	ACTAC	GGAG	CCAACTTT	CCAA	CCAAAT	GGTAAT	TTTAT	TGCGGT	CTATG	960
Qy	961	GACTG	TTG	AC	CCCTCT	TGT	CGAAT	TATAT	TG	TCG	ACAGT	1020
Db	961	GACTG	TTG	AC	CCCTCT	TGT	CGAAT	TATAT	TG	TCG	ACAGT	1020
Qy	1021	AGG	AGCA	AC	CGCTA	AGG	GGAC	CCATCT	GT	TGAT	G	1080
Db	1021	AGG	AGCA	AC	CGCTA	AGG	GGAC	CCATCT	GT	TGAT	G	1080
Qy	1081	TCTT	AG	GTCA	ATCA	CCCT	CCATTA	AGG	GA	TTGCC	ACA	1140
Db	1081	TCTT	AG	GTCA	ATCA	CCCT	CCATTA	AGG	GA	TTGCC	ACA	1140
Qy	1141	TCG	AG	ATCG	AAAC	CG	ACG	AGTGC	AGATTT	CTG	T	1200
Db	1141	TCG	AG	ATCG	AAAC	CG	ACG	AGTGC	AGATTT	CTG	T	1200
Qy	1201	AAA	CTTTAG	GGAT	GAATAT	TGGG	AAAAAT	GTATG	AAAGT	CGCG	CTTACT	1260
Db	1201	AAA	CTTTAG	GGAT	GAATAT	TGGG	AAAAAT	GTATG	AAAGT	CGCG	CTTACT	1260

Db	1201	AACTTAGGATGAATATGGGAAATGATGAAGTCGCCCTTACTGTAGAGGCTATCA	1260
Qy <th>1261</th> <th>AAGTAGCGGAAGTGCCTAAATCTATATAGCAATACACATAAGAAATTAACCGTAAACCCCTCTCTC</th> <th>1320</th>	1261	AAGTAGCGGAAGTGCCTAAATCTATATAGCAATACACATAAGAAATTAACCGTAAACCCCTCTCTC	1320
Db <th>1261</th> <th>AAGTAGCGGAAGTGCCTAAATCTATATAGCAATACACATAAGAAATTAACCGTAAACCCCTCTCTC</th> <th>1320</th>	1261	AAGTAGCGGAAGTGCCTAAATCTATATAGCAATACACATAAGAAATTAACCGTAAACCCCTCTCTC	1320
Qy <th>1321</th> <th>AACATTAGTAATGACGAGAGCAATACTTTGGGATAAAAAAACAATTAATAATCCTTATCTCT</th> <th>1380</th>	1321	AACATTAGTAATGACGAGAGCAATACTTTGGGATAAAAAAACAATTAATAATCCTTATCTCT	1380
Db <th>1321</th> <th>AACATTAGTAATGACGAGAGCAATACTTTGGGATAAAAAAACAATTAATAATCCTTATCTCT</th> <th>1380</th>	1321	AACATTAGTAATGACGAGAGCAATACTTTGGGATAAAAAAACAATTAATAATCCTTATCTCT	1380
Qy <th>1381</th> <th>TTCCGGTTCAGTTCTCATATATTTTCAATATACCTCCCGTTGGATCTTTTCCAAAGGGGAGG</th> <th>1440</th>	1381	TTCCGGTTCAGTTCTCATATATTTTCAATATACCTCCCGTTGGATCTTTTCCAAAGGGGAGG	1440
Db <th>1381</th> <th>TTCCGGTTCAGTTCTCATATATTTTCAATATACCTCCCGTTGGATCTTTTCCAAAGGGGAGG</th> <th>1440</th>	1381	TTCCGGTTCAGTTCTCATATATTTTCAATATACCTCCCGTTGGATCTTTTCCAAAGGGGAGG	1440
Qy <th>1441</th> <th>TTTTATTGGAAAGGTTAAGTATAGTATACATCCGATTCATCCAGAGGAATGCTTTGAACA</th> <th>1500</th>	1441	TTTTATTGGAAAGGTTAAGTATAGTATACATCCGATTCATCCAGAGGAATGCTTTGAACA	1500
Db <th>1441</th> <th>TTTTATTGGAAAGGTTAAGTATAGTATACATCCGATTCATCCAGAGGAATGCTTTGAACA</th> <th>1500</th>	1441	TTTTATTGGAAAGGTTAAGTATAGTATACATCCGATTCATCCAGAGGAATGCTTTGAACA	1500
Qy <th>1501</th> <th>CCTCCGTCACCTAG 1513</th> <th></th>	1501	CCTCCGTCACCTAG 1513	
Db <th>1501</th> <th>CCTCCGTCACCTAG 1513</th> <th></th>	1501	CCTCCGTCACCTAG 1513	
RESULT 2			
A48232			
LOCUS			
DEFINITION	Sequence 11 from Patent EP0698667.		
ACCESSION	A48232		
VERSION	A48232.1 GI:2302079		
KEYWORDS	unidentified		
SOURCE	unidentified		
ORGANISM	unclassified		
REFERENCE	1 (bases 1 to 1513)		
AUTHORS	De,B.E., Lahaye,A., Ledoux,P. and Detroz,R.		
TITLE	Xylanase, microorganisms for its production, DNA molecules, processes of preparation and use thereof		
JOURNAL	Patent: EP 0698667-A 11 28-FEB-1996;		
SOLVAY (BE)			
COMMENT	Other publication BE 1008751 960702		
	Other publication BE 1008570 960604		
	Other publication BR 9503454 960305		
	Other publication JP 8092284 960409		
	Other publication FI 953578 960127		
	Other publication CA 2154628 960127		
FEATURES	Other publication AU 2508695 960208.		
	Location/Qualifiers		
source	1..1513		
	/organism="unidentified"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32644"		
	620..1366		
	/note="unnamed protein product"		
CDS	/codon_start=1		
	/protein_id="CAA03094.1"		
	/db_xref="GI:2302080"		
	/translation="MRCKLTLLAFLVCFALTLPASIIQAIIVTDSIGNHGDYDEVYEFV		
	FKWDSGSGFMILNHGTFSAQWNNVNNILFRKGFKEFTQHOVCNMSINVCANFC		
	PNGNAYLCVIGWTVDPLEVEYIIVDSWGNRPPGATPKGTIVDGGTDIVETLWVQF		
sig_peptide	SIKGIATFKQVSVRRSKYSGTISVNSNFRPMENLGNMGMKMTVEALTVEGYOSSGS		
	ANVYSNLTIRNGNPLSTISNDESLTKNN"		
	620..700		
	mat_peptide		
ORIGIN	701..1363		
	/product="unnamed"		
	Query Match		
	Best Local Similarity 100.0%; Score 1513; DB 6; Length 1513;		
Matches 1513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy <th>1</th> <th>AAATTGAATGTGTATATCTATATGATTAACGACAATCGTCACTCTTTTAACCTAATCTC</th> <th>60</th>	1	AAATTGAATGTGTATATCTATATGATTAACGACAATCGTCACTCTTTTAACCTAATCTC	60

```
Db 1 AAATTGAATTGTGTATCTAATGATAACGACAAATCGTCACTGTTTTTAAACTAATCTC 60
Qy 61 AACCAATCTCTTTAATTAACGTAACCACTTGCATCTTTATCACAGACATCTTT 120
Db 61 AACCAATCTCTTTAATTAACGTAACCACTTGCATCTTTATCACAGACATCTTT 120
Qy 121 ATAGGAACCTTCCCATTTGCAAGACGATAAAAACTTTTCCCTATTTATCTTATCG 180
Db 121 ATAGGAACCTTCCCATTTGCAAGACGATAAAAACTTTTCCCTATTTATCTTATCG 180
Qy 181 CCTTGATCGGTTAAATTTGTAACCTTTATTTAGTTTACGTGATGTTCCCTCATTCATC 240
Db 181 CCTTGATCGGTTAAATTTGTAACCTTTATTTAGTTTACGTGATGTTCCCTCATTCATC 240
Qy 241 CATTAATCACAGTTAAGCTAGAGTCACTCTTTTTCGGTTCCTCAAAATACCTGAGAAC 300
Db 241 CATTAATCACAGTTAAGCTAGAGTCACTCTTTTTCGGTTCCTCAAAATACCTGAGAAC 300
Qy 301 ATTTATGTCATATTTTCTCAGCGCGCTCCATATGGAATATATATATCTTTTATACATA 360
Db 301 ATTTATGTCATATTTTCTCAGCGCGCTCCATATGGAATATATATATCTTTTATACATA 360
Qy 361 TTAAGTAAATTTAGTATATATCTTGCCTTATCAAAATGTGAGATAATCTAATTGATCAACA 420
Db 361 TTAAGTAAATTTAGTATATATCTTGCCTTATCAAAATGTGAGATAATCTAATTGATCAACA 420
Qy 421 AGCAGCTATCCAAAAAAGCTGATGTTGACCTTTAAAGAAAGTGCCTATCTATCAAAA 480
Db 421 AGCAGCTATCCAAAAAAGCTGATGTTGACCTTTAAAGAAAGTGCCTATCTATCAAAA 480
Qy 481 GATAATTTCCAGTTTCAAAATTTGAAATTTAGTGTGATGGAATAGTTGAATGTCAACTG 540
Db 481 GATAATTTCCAGTTTCAAAATTTGAAATTTAGTGTGATGGAATAGTTGAATGTCAACTG 540
Qy 541 CTGTGAAAGAGGAGGTAGTAGTACCGTAGACTTTCATTACCAAAAATTTAGTTGTAACAAA 600
Db 541 CTGTGAAAGAGGAGGTAGTAGTACCGTAGACTTTCATTACCAAAAATTTAGTTGTAACAAA 600
Qy 601 TTAAGAGGAGGATGCTTAATGAGACAAAGAAATTTGACGTGATTTAGCTTTTACT 660
Db 601 TTAAGAGGAGGATGCTTAATGAGACAAAGAAATTTGACGTGATTTAGCTTTTACT 660
Qy 661 TTGTTTGTCACTAACCTTACCTCAGAAATAATTTAGGCAAAATCGTACCACCAATTC 720
Db 661 TTGTTTGTCACTAACCTTACCTCAGAAATAATTTAGGCAAAATCGTACCACCAATTC 720
Qy 721 CATTTGCAACACGATGGCTATGATTAATGAATTTTGGAAAGATAGCGGTGGCTCTGGAC 780
Db 721 CATTTGCAACACGATGGCTATGATTAATGAATTTTGGAAAGATAGCGGTGGCTCTGGAC 780
Qy 781 AATGATTTCTCAATCATGGCGGTAGCTTCACTGCGCCCAATGGAACAAATGTTAACAATATT 840
Db 781 AATGATTTCTCAATCATGGCGGTAGCTTCACTGCGCCCAATGGAACAAATGTTAACAATATT 840
Qy 841 ATTCGGTAAAGGTAAATAATTTCAATGAACACAAACACACAAAGTTCGTTAACAATGTC 900
Db 841 ATTCGGTAAAGGTAAATAATTTCAATGAACACAAACACACAAAGTTCGTTAACAATGTC 900
Qy 901 CATAAACTACGAGCCAACTTCCAAACCAATGTTAATTCGCTATTTATGCGTCTATGGTTG 960
Db 901 CATAAACTACGAGCCAACTTCCAAACCAATGTTAATTCGCTATTTATGCGTCTATGGTTG 960
Qy 961 GACTGTTGACCTCTGTGGAATATATATTTGTCGACAGTTTGGGGCAACTGGCGTCCACC 1020
Db 961 GACTGTTGACCTCTGTGGAATATATATTTGTCGACAGTTTGGGGCAACTGGCGTCCACC 1020
Qy 1021 AGGAGCAACGCTTAAGGGACCATCACTGTTGATGGAGGACATATCATATCTACAGAC 1080
Db 1021 AGGAGCAACGCTTAAGGGACCATCACTGTTGATGGAGGACATATCATATCTACAGAC 1080
Qy 1081 TCTTAGAGTCAATCAACCTCCATTAAAGGGGATTTGCCACATTTTAAACAAATTTGAGTGT 1140
Db 1081 TCTTAGAGTCAATCAACCTCCATTAAAGGGGATTTGCCACATTTTAAACAAATTTGAGTGT 1140
```

```
Qy 1141 TCGAAGATCGAAACGACGAGTGGCAGCAATTTCTGTCTAGCAACCACTTTAGAGCGTGGGA 1200
Db 1141 TCGAAGATCGAAACGACGAGTGGCAGCAATTTCTGTCTAGCAACCACTTTAGAGCGTGGGA 1200
Qy 1201 AAACCTTAGGATGAATATGGGAAAATGTATGAAGTCGGCTTACTGTAGAGGCTATCA 1260
Db 1201 AAACCTTAGGATGAATATGGGAAAATGTATGAAGTCGGCTTACTGTAGAGGCTATCA 1260
Qy 1261 AAGTAGCGGAAGTGTAAATGTATATAGCAATACACTAAGAAATTAACGGTAACCTCTCTC 1320
Db 1261 AAGTAGCGGAAGTGTAAATGTATATAGCAATACACTAAGAAATTAACGGTAACCTCTCTC 1320
Qy 1321 AACTATTAGTAATGACGAGACATAAATTTGGATAAAAACAATTAATAATTCCTTATCTCT 1380
Db 1321 AACTATTAGTAATGACGAGACATAAATTTGGATAAAAACAATTAATAATTCCTTATCTCT 1380
Qy 1381 TTCCGTTCCAGTTCTCATTTTCAATAAATCCCGGTGGATCTTTTCCAAACGGGAGG 1440
Db 1381 TTCCGTTCCAGTTCTCATTTTCAATAAATCCCGGTGGATCTTTTCCAAACGGGAGG 1440
Qy 1441 TTTTATTGGAAGGTTAAGTATAGTATCTCCGATTCCATCCAGAGGAATGCTTGAACA 1500
Db 1441 TTTTATTGGAAGGTTAAGTATAGTATCTCCGATTCCATCCAGAGGAATGCTTGAACA 1500
Qy 1501 CCTCCGTCACCTAG 1513
Db 1501 CCTCCGTCACCTAG 1513
```

## RESULT 3

```
AR193055 LOCUS 1513 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 10 from patent US 6346407.
ACCESSION AR193055
VERSION AR193055.1 GI:20239020
KEYWORDS Unknow.
SOURCE Unknow.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1513)
AUTHORS De Buy, E., Lahaye, A., Ledoux, P. and Detroz, R.
TITLE Xylanase, microorganisms producing it, DNA molecules, methods for
preparing this xylanase and uses of the latter
JOURNAL Patent: US 6346407-A 10 12-FEB-2002;
FEATURES Location/Qualifiers
source 1. .1513
/organism="unknown"
/mol_type="unassigned DNA"
```

## ORIGIN

```
Query Match 100.0%; Score 1513; DB 6; Length 1513;
Best Local Similarity 100.0%; Pred. No. 1.5e-295; Indels 0; Gaps 0;
Matches 1513; Conservative 0; Mismatches 0;

Qy 1 AAATTGAATTGTGTATCTAATGATAACGACAAATCGTCACTGTTTTTAAACTAATCTC 60
Db 1 AAATTGAATTGTGTATCTAATGATAACGACAAATCGTCACTGTTTTTAAACTAATCTC 60
Qy 61 AACCAATCTCTTTAATTAACGTAACCACTTGCATCTTTATCACAGACATCTTT 120
Db 61 AACCAATCTCTTTAATTAACGTAACCACTTGCATCTTTATCACAGACATCTTT 120
Qy 121 ATAGGAACCTTCCCATTTGCAAGACGATAAAAACTTTTCCCTATTTATCTTATCG 180
Db 121 ATAGGAACCTTCCCATTTGCAAGACGATAAAAACTTTTCCCTATTTATCTTATCG 180
Qy 181 CCTTGATCGGTTAAATTTGTAACCTTTATTTAGTTTACGTGATGTTCCCTCATTCATC 240
Db 181 CCTTGATCGGTTAAATTTGTAACCTTTATTTAGTTTACGTGATGTTCCCTCATTCATC 240
Qy 241 CATTAATCACAGTTAAGCTAGAGTCACTCTTTTTCGGTTCCTCAAAATACCTGAGAAC 300
Db 241 CATTAATCACAGTTAAGCTAGAGTCACTCTTTTTCGGTTCCTCAAAATACCTGAGAAC 300
```

241 CATTAATACAGTTAAACGCTAGAGTCATCTTTTTCGGTTCTCAAAATACCTGAGAAC 300  
 301 ATTTATGTCATATTTTCTCAGCGGCTCCATATGGAATATATATCTTTTATACATA 360  
 301 ATTTATGTCATATTTTCTCAGCGGCTCCATATGGAATATATATCTTTTATACATA 360  
 361 TTAAGTAAATAGTATATATCTTGGGTTATCAAAATGTGAGATACTTAATTCATCAACA 420  
 361 TTAAGTAAATAGTATATATCTTGGGTTATCAAAATGTGAGATACTTAATTCATCAACA 420  
 421 AGCAGCTATCCAAAAAACAACCTGATGTGACCTCTTAAAGAAAGTGCATCTATCTAGAAA 480  
 421 AGCAGCTATCCAAAAAACAACCTGATGTGACCTCTTAAAGAAAGTGCATCTATCTAGAAA 480  
 481 GATAATATCCAGTTTCAAAATTTGAAATAGTGTGTATGGAATAGTTGAATGTCAACTG 540  
 481 GATAATATCCAGTTTCAAAATTTGAAATAGTGTGTATGGAATAGTTGAATGTCAACTG 540  
 541 CTGTGAAAGGAGGTAGGTAGTACCGTAGACTTCAATACCAAAATTTAGTTGTAATAAAA 600  
 541 CTGTGAAAGGAGGTAGGTAGTACCGTAGACTTCAATACCAAAATTTAGTTGTAATAAAA 600  
 601 TTAAGAGGAGTAATGCTTAATGAGACAAAGAAATTTGACGTTTATAGCTTTTATAGT 660  
 601 TTAAGAGGAGTAATGCTTAATGAGACAAAGAAATTTGACGTTTATAGCTTTTATAGT 660  
 661 TTGTTTTCACCTAACCTTACCTGAGAAATATTCAGGACAAATTCGTCACCGCAATTC 720  
 661 TTGTTTTCACCTAACCTTACCTGAGAAATATTCAGGACAAATTCGTCACCGCAATTC 720  
 721 CATTTGGCAACACGATGCTATGATTTATGAAATTTTGGAAAGATAGCGGTGCTCTGGAC 780  
 721 CATTTGGCAACACGATGCTATGATTTATGAAATTTTGGAAAGATAGCGGTGCTCTGGAC 780  
 781 AATGATTTCAATCATGCGGTAAGTTCAGTCCCAATGGAACAATTTAAACAATAT 840  
 781 AATGATTTCAATCATGCGGTAAGTTCAGTCCCAATGGAACAATTTAAACAATAT 840  
 841 ATTCCGTAAGGTAAATAATTCAAATGAAACACACACACCAACCAAGTTGGTAACATGC 900  
 841 ATTCCGTAAGGTAAATAATTCAAATGAAACACACACACCAACCAAGTTGGTAACATGC 900  
 901 CATAAACAACGAGGCAACTTCCAAACCAATGTAATGCTATTTATGCTCTATGGTTG 960  
 901 CATAAACAACGAGGCAACTTCCAAACCAATGTAATGCTATTTATGCTCTATGGTTG 960  
 961 GACTGTGACCTCTTGTGCAATATATTTGTCAGACAGTTGGGCAACCTGGCGTCCACC 1020  
 961 GACTGTGACCTCTTGTGCAATATATTTGTCAGACAGTTGGGCAACCTGGCGTCCACC 1020  
 1021 AGGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGAACATATGATATCTACGAGAC 1080  
 1021 AGGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGAACATATGATATCTACGAGAC 1080  
 1081 TCTTAGAGTCAATCAACCTTCCATTAAGGGGATTCGCCATTTAAACAATATGGAGTGT 1140  
 1081 TCTTAGAGTCAATCAACCTTCCATTAAGGGGATTCGCCATTTAAACAATATGGAGTGT 1140  
 1141 TCGAAGATCGAAACGACGAGTGGCAAGATTTCTGTGACACACCACTTTAGAGCGTGGGA 1200  
 1141 TCGAAGATCGAAACGACGAGTGGCAAGATTTCTGTGACACACCACTTTAGAGCGTGGGA 1200  
 1201 AAACCTTAGGATGAATATGGGGAATATGATGAAGTCCGCTTACTGTAGAGGCTATCA 1260  
 1201 AAACCTTAGGATGAATATGGGGAATATGATGAAGTCCGCTTACTGTAGAGGCTATCA 1260  
 1261 AAGTAGCGGAAGTGTATATAGTATATAGCAATCACTAAAGAATTAACGGTAACCCCTCTCTC 1320  
 1261 AAGTAGCGGAAGTGTATATAGTATATAGCAATCACTAAAGAATTAACGGTAACCCCTCTCTC 1320  
 1321 AACTATTAGTATGACGAGCATAACTTGGATAAAAACAATTAATAAATCCTTATCTCT 1380  
 1321 AACTATTAGTATGACGAGCATAACTTGGATAAAAACAATTAATAAATCCTTATCTCT 1380

1381 TTCGGTTCAGTTCTCATTTATTTTCAATAACCTCCCGTTGCAATCTTTTCAACGGGAGG 1440  
 1381 TTCGGTTCAGTTCTCATTTATTTTCAATAACCTCCCGTTGCAATCTTTTCAACGGGAGG 1440  
 1441 TTTTATTGGAAGGTTAAGTATAGTATCTCCGATTCATCCAGAGGAATGCTTGAACA 1500  
 1441 TTTTATTGGAAGGTTAAGTATAGTATCTCCGATTCATCCAGAGGAATGCTTGAACA 1500  
 1501 CCTCGTCACTAG 1513  
 1501 CCTCGTCACTAG 1513

RESULT 4  
 AR193056  
 LOCUS 1513 bp DNA linear PAT 20-APR-2002  
 DEFINITION Sequence 11 from patent US 6346407.  
 ACCESSION AR193056  
 VERSION AR193056.1 GI:20239021  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1513)  
 AUTHORS De Buy, E., Lahaye, A., Ledoux, P. and Detroz, R.  
 TITLE Xylanase, microorganisms producing it, DNA molecules, methods for preparing this xylanase and uses of the latter  
 JOURNAL Patent: US 6346407-A 11 12-FEB-2002;  
 FEATURES  
 source  
 1. 1513  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 1513; DB 6; Length 1513;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-295;  
 Matches 1513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATGGAATGTGTATATCTTAATGATAACGACAATCGTCACTGTTTAACTAATCTC 60  
 DB 1 AAATGGAATGTGTATATCTTAATGATAACGACAATCGTCACTGTTTAACTAATCTC 60  
 QY 61 AAACCAATACCTCTTTTAAACGCTAACCACTTTCATCTTATACAAAGAACATCTTT 120  
 DB 61 AAACCAATACCTCTTTTAAACGCTAACCACTTTCATCTTATACAAAGAACATCTTT 120  
 QY 121 ATAGGAACCTTCCCATTTGCAAGACGATAAAAAATCTTTTCCCTATTTTATCTTATCG 180  
 DB 121 ATAGGAACCTTCCCATTTGCAAGACGATAAAAAATCTTTTCCCTATTTTATCTTATCG 180  
 QY 181 CTTGATCGGTTTAAATTTGTAACCTTTTATTTAGTTTACGTGATGTTCCCTCATCTAC 240  
 DB 181 CTTGATCGGTTTAAATTTGTAACCTTTTATTTAGTTTACGTGATGTTCCCTCATCTAC 240  
 QY 241 CATTAATCACAGTTTAAACGCTAGAGTCACTTTTTCGGTCTCAAAATACCTCAAGAAC 300  
 DB 241 CATTAATCACAGTTTAAACGCTAGAGTCACTTTTTCGGTCTCAAAATACCTCAAGAAC 300  
 QY 301 ATTTATGTCATATTTTCTCAGCGGCTCCATATGGAATATATATCTTTTATACATA 360  
 DB 301 ATTTATGTCATATTTTCTCAGCGGCTCCATATGGAATATATATCTTTTATACATA 360  
 QY 361 TTAAGTAAATAGTATATCTTGGTTCATCAAAATGAGATATCTTAATTTGATCAACA 420  
 DB 361 TTAAGTAAATAGTATATCTTGGTTCATCAAAATGAGATATCTTAATTTGATCAACA 420  
 QY 421 AGCAGCTATCCAAAAAACAACCTGATGTGACCTCTTAAAGAAAGTGTCACTATCTATGAAA 480  
 DB 421 AGCAGCTATCCAAAAAACAACCTGATGTGACCTCTTAAAGAAAGTGTCACTATCTATGAAA 480  
 QY 481 GATAATATCCAGTTTCAAAATTTGAAATAGTGTGTATGGAATAGTTGAATGTCAACTG 540  
 DB 481 GATAATATCCAGTTTCAAAATTTGAAATAGTGTGTATGGAATAGTTGAATGTCAACTG 540

Db	481	GATTAATATCCAGTTTCAAAATTTGAATAGTGTGTATGAATAGTTTCAATGTCACTG	540
Qy	541	CTGTGAAGAGGGTAGTACCGTAGACTTCAATTAACCAAAATAGTTGTAATAAAA	600
Db	541	CTGTGAAGAGGGTAGTACCGTAGACTTCAATTAACCAAAATAGTTGTAATAAAA	600
Qy	601	TTAAAAGGGAATGCTTAATGACCAAAAGAAATGAGTTGATTTTAGCCCTTTTAGT	660
Db	601	TTAAAAGGGAATGCTTAATGACCAAAAGAAATGAGTTGATTTTAGCCCTTTTAGT	660
Qy	661	TTGTTTTCACCTAACCTTACCTGCGAGAAATATTCAGGCACAAATCGTCACCGACAATTC	720
Db	661	TTGTTTTCACCTAACCTTACCTGCGAGAAATATTCAGGCACAAATCGTCACCGACAATTC	720
Qy	721	CATTGGCAACACGATGGCTATGATTAATGAATTTTGGAAAGATAGCGGTGGCTCTGGAC	780
Db	721	CATTGGCAACACGATGGCTATGATTAATGAATTTTGGAAAGATAGCGGTGGCTCTGGAC	780
Qy	781	AATGATTCATCATGGCGGTACGTTTCACTGCCCCAATGGAACAATGTTAACACAATTC	840
Db	781	AATGATTCATCATGGCGGTACGTTTCACTGCCCCAATGGAACAATGTTAACACAATTC	840
Qy	841	ATTCCGTAAGGTAAATAATCAATGAAACACAAACACACCAACCAAGTTGGTAACATGTC	900
Db	841	ATTCCGTAAGGTAAATAATCAATGAAACACAAACACACCAACCAAGTTGGTAACATGTC	900
Qy	901	CATAAATACGAGGACCACTTCCAAACCAATGGTAATGGTATTTAGCGTCTATGGTTG	960
Db	901	CATAAATACGAGGACCACTTCCAAACCAATGGTAATGGTATTTAGCGTCTATGGTTG	960
Qy	961	GACTGTGACCTCTTGTGCAATATTTATGTCGACAGTTGGGCAACTGGCGTCCACC	1020
Db	961	GACTGTGACCTCTTGTGCAATATTTATGTCGACAGTTGGGCAACTGGCGTCCACC	1020
Qy	1021	AGGAGCAACCCCTAAGGGACCACTCACTGTTGATGGAGAAACATATGATATCTACGAGAC	1080
Db	1021	AGGAGCAACCCCTAAGGGACCACTCACTGTTGATGGAGAAACATATGATATCTACGAGAC	1080
Qy	1081	TCCTAGAGTCAATCAACCCCTCCATTAAGGGATGGCCACATTTAAACAATATTTGGAGTGT	1140
Db	1081	TCCTAGAGTCAATCAACCCCTCCATTAAGGGATGGCCACATTTAAACAATATTTGGAGTGT	1140
Qy	1141	TCGAAGTCGAACGACGAGTGGCAGATTTCTGTGACCAACCTTTTAGACGCTATCA	1200
Db	1141	TCGAAGTCGAACGACGAGTGGCAGATTTCTGTGACCAACCTTTTAGACGCTATCA	1200
Qy	1201	AAACTTAGGATGAATATGGGAAATGTATGAAGTCGCGCTTACTGTAGAAGCTATCA	1260
Db	1201	AAACTTAGGATGAATATGGGAAATGTATGAAGTCGCGCTTACTGTAGAAGCTATCA	1260
Qy	1261	AAGTAGCGGAAGTCTAATGTATATAGCAATACATTAAGAAATTAACGGTAACCTCTCTC	1320
Db	1261	AAGTAGCGGAAGTCTAATGTATATAGCAATACATTAAGAAATTAACGGTAACCTCTCTC	1320
Qy	1321	AACATTTAGTAAGACGAGCATACTTTGGATTAATAAACAATTTAAATTCCTTACTCT	1380
Db	1321	AACATTTAGTAAGACGAGCATACTTTGGATTAATAAACAATTTAAATTCCTTACTCT	1380
Qy	1381	TTGGTTTCAGTCTCATTTATTTTCAAAATAACCTCCCGTTGGATCTTTTCCACGGGAGG	1440
Db	1381	TTGGTTTCAGTCTCATTTATTTTCAAAATAACCTCCCGTTGGATCTTTTCCACGGGAGG	1440
Qy	1441	TTTTTTTGAAGGTTAAGTATAGTATATCTCCGATTCATCCAGAGAAATGCTTGAACA	1500
Db	1441	TTTTTTTGAAGGTTAAGTATAGTATATCTCCGATTCATCCAGAGAAATGCTTGAACA	1500
Qy	1501	CTCCGTCCTAG	1513
Db	1501	CTCCGTCCTAG	1513
RESULT 5			
A48225			

LOCUS	A48225	744 bp	DNA	linear	PAT 07-MAR-1997			
DEFINITION	Sequence 4 from Patent EP0698667.							
ACCESSION	A48225							
VERSION	A48225.1	GI:2302072						
KEYWORDS	unidentified							
SOURCE	unclassified							
ORGANISM	1 (bases 1 to 744)							
REFERENCE	De,B.E., Lahaye,A., Ledoux,P. and Detroz,R.							
AUTHORS	Xylanase, microorganisms for its production, DNA molecules, process							
TITLE	of preparation and use thereof							
JOURNAL	Patent: EP 0698667-A 4 28-FEB-1996;							
	SOLVAY (BE)							
COMMENT	Other publication BE 1008751 960702							
	Other publication BE 1008570 960604							
	Other publication BR 9503454 960305							
	Other publication JP 8092284 960409							
	Other publication FI 953578 960127							
	Other publication CA 2154628 960127							
	Other publication AU 2508695 960208.							
FEATURES	Location/Qualifiers							
source	1..744							
	/organism="unidentified"							
	/mol_type="unassigned DNA"							
	/db_xref="taxon:32644"							
ORIGIN								
Query Match	49.2%; Score 744; DB 6; Length 744;							
Best Local Similarity	100.0%; Pred. No. 3.9e-140;							
Matches	744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
Qy	620	ATGAGCAAAAGAAATTCAGCTTTCATTTTAGCCCTTTTGTAGTTTGTTCACCTTA	679					
Db	1	ATGAGCAAAAGAAATTCAGCTTTCATTTTAGCCCTTTTGTAGTTTGTTCACCTTA	60					
Qy	680	CCTGCAGAAATTAATTCAGGCACAAATCGTCACGCAATTCATTCGGCAACACGATGC	739					
Db	61	CCTGCAGAAATTAATTCAGGCACAAATCGTCACGCAATTCATTCGGCAACACGATGC	120					
Qy	740	TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTGGGCAATGATTTCTCAATCAGC	799					
Db	121	TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTGGGCAATGATTTCTCAATCAGC	180					
Qy	800	GGTACGTTTCAGTGCCTCAATGGAAACAAATGTTAAACAATATTTCCGTAAGGTAAAAA	859					
Db	181	GGTACGTTTCAGTGCCTCAATGGAAACAAATGTTAAACAATATTTCCGTAAGGTAAAAA	240					
Qy	860	TTCAATGAAACACAAACACACCAACCAAGTTGGTAAACATGTCCATAAACTACGGAGCCAC	919					
Db	241	TTCAATGAAACACAAACACACCAACCAAGTTGGTAAACATGTCCATAAACTACGGAGCCAC	300					
Qy	920	TTCCAAACCAATGGTAAATGGTATTTATGCGTCTATGGTGGACTGTGGACCTCTTGTC	979					
Db	301	TTCCAAACCAATGGTAAATGGTATTTATGCGTCTATGGTGGACTGTGGACCTCTTGTC	360					
Qy	980	GAATATTTATTTGTCGACAGTTGGGCAACTGGCGTCCACAGAGCAACGCTTAAGGGG	1039					
Db	361	GAATATTTATTTGTCGACAGTTGGGCAACTGGCGTCCACAGAGCAACGCTTAAGGGG	420					
Qy	1040	ACCATCACTGTTGATGGAGGAAACATATGATATCTACGAGCTCTTAGAGTCAATCAACCC	1099					
Db	421	ACCATCACTGTTGATGGAGGAAACATATGATATCTACGAGCTCTTAGAGTCAATCAACCC	480					
Qy	1100	TCCATTTAAGGGGATTCGCACATTTAAACAATTTGGAGTGTTCGAAGATCGAAACGCACG	1159					
Db	481	TCCATTTAAGGGGATTCGCACATTTAAACAATTTGGAGTGTTCGAAGATCGAAACGCACG	540					
Qy	1160	AGTGGCACGATTTCTGTGACGAAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG	1219					
Db	541	AGTGGCACGATTTCTGTGACGAAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATG	600					
Qy	1220	GGGAAATGTATGAAGTCCGCGTCTACTGTAGAAGGCTATCAAAGTAGCGGAGTCTAAT	1279					



Db 601 GGGAAAAATGTAGAAAGTCGCGCTTACTGTAGAGGCTATCAAAGTAGCGGAAGTCTTAAT 660  
Qy 1280 GTATATAGCAATACACTAAGAAATTAACGTTACCCCTCTCTCAACTATTAGTAATGACGAG 1339  
Db 661 GTATATAGCAATACACTAAGAAATTAACGTTACCCCTCTCTCAACTATTAGTAATGACGAG 720  
Qy 1340 AGCATAACTTTGGATAAAAAAACAAT 1363  
Db 721 AGCATAACTTTGGATAAAAAAACAAT 744  
RESULT 6  
LOCUS A48226 744 bp DNA linear PAT 07-MAR-1997  
DEFINITION Sequence 5 from Patent EP0698667.  
ACCESSION A48226  
VERSION A48226.1 GI:2302073  
KEYWORDS unidentified  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 744)  
AUTHORS De,B.E., Lahaye,A., Ledoux,P. and Detroz,R.  
TITLE Xylanase, microorganisms for its production, DNA molecules, process of preparation and use thereof  
JOURNAL Patent: EP 0698667-A 5 28-FEB-1996;  
SOLVAY (BE)  
COMMENT Other publication BE 1008751 960702  
Other publication BE 1008570 960604  
Other publication BR 9503454 960305  
Other publication JP 8092284 960409  
Other publication FI 953578 960127  
Other publication CA 2154628 960127  
Other publication AU 2508695 960208.  
FEATURES  
source  
1..744  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
CDS  
1..5744  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAA03092.1"  
/db\_xref="GI:2302074"  
/translation="MRQKLLTLLAFVCFALTPAEIIQAIQIVTDSIGNHGDYDE  
FWKSGSGTMLNHHGGTFSQNNNNILFRKKFNETQHQGVNMSINYGNFQ  
PNGNAYLCVGMVDPVLEYYIVDSWGNRPFGATPKGTITVDGGYDIYETLRVNP  
SIKGIATFKQYMSVRSRSTSGTISVSNHFRAWENLGNMNGMKYVEALTVEGYQSSGS  
ANVYSNTRLINGNPLSTISNDESITLDKNN"  
sig\_peptide 1..81  
mat\_peptide 82..744  
/product="unnamed"  
ORIGIN  
Query Match 49.2%; Score 744; DB 6; Length 744;  
Best Local Similarity 100.0%; Pred. No. 3.9e-140; Mismatches 0; Indels 0; Gaps 0;  
Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 620 ATGAGACAAAAGAAATGACGTTGATTTTAGCCTTTTATAGCTTTTATAGTTTGTGCACTAACCTTA 679  
Db 1 ATGAGACAAAAGAAATGACGTTGATTTTAGCCTTTTATAGCTTTTATAGTTTGTGCACTAACCTTA 60  
Qy 680 CCTGAGAAATAATTCAGGCACAAATCGTCACCGCAATTCATTCGGCAACACGATGCG 739  
Db 61 CCTGAGAAATAATTCAGGCACAAATCGTCACCGCAATTCATTCGGCAACACGATGCG 120  
Qy 740 TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTCGGACAATGATTCTCAATCATGCG 799  
Db 121 TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTCGGACAATGATTCTCAATCATGCG 180  
Qy 800 GGTACGTTACGTCGCAATGGAACAATGTTTAAACAATATTATTCGTTAAAGGTTAAAAA 859

Db 181 GGTACGTTACGTCGCAATGGAACAATGTTTAAACAATATTATTCCGTTAAAGGTTAAAAA 240  
Qy 860 TTCAATGAAACACAAAACACACAACTTGTGTAACATGTCCATAAACTACGAGCCCAAC 919  
Db 241 TTCAATGAAACACAAAACACACAACTTGTGTAACATGTCCATAAACTACGAGCCCAAC 300  
Qy 920 TTCCAAACCAATGTTGTAATGCGTATTATGCGTCTATGCGTCTGTTGGACTGTTTGGACCTCTTGTG 979  
Db 301 TTCCAAACCAATGTTGTAATGCGTATTATGCGTCTGTTGGACTGTTTGGACCTCTTGTG 360  
Qy 980 GAATATTATATGTCGACAGTTGGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGGG 1039  
Db 361 GAATATTATATGTCGACAGTTGGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGGG 420  
Qy 1040 ACCATCACTGTTGATGGAGGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 1099  
Db 421 ACCATCACTGTTGATGGAGGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480  
Qy 1100 TCCATTAAGGGGATTGCCACATTTAAACAATATTGGAGTGTTCGAAGATCGAAACGCAAG 1159  
Db 481 TCCATTAAGGGGATTGCCACATTTAAACAATATTGGAGTGTTCGAAGATCGAAACGCAAG 540  
Qy 1160 AGTGGCAGATTTCTGTCAGCAACACTTTTAGAGGTGGGAAACTTAGGGATGAATATG 1219  
Db 541 AGTGGCAGATTTCTGTCAGCAACCACTTTTAGAGGTGGGAAACTTAGGGATGAATATG 600  
Qy 1220 GGGAAATGATGAAGTCGCGCTTACTGTAGAAGCTATCAAGTAGCGGAAGTGTCTAAT 1279  
Db 601 GGGAAATGATGAAGTCGCGCTTACTGTAGAAGCTATCAAGTAGCGGAAGTGTCTAAT 660  
Qy 1280 GTATATAGCAATACACTAAGAAATTAACGTTAAACCTCTCTCAACTATTAGTAATGACGAG 1339  
Db 661 GTATATAGCAATACACTAAGAAATTAACGTTAAACCTCTCTCAACTATTAGTAATGACGAG 720  
Qy 1340 AGCATACTTTGGATAAAAAACAAT 1363  
Db 721 AGCATACTTTGGATAAAAAACAAT 744  
RESULT 7  
LOCUS AR193051 744 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 4 from patent US 6346407.  
ACCESSION AR193051  
VERSION AR193051.1 GI:20239016  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 744)  
AUTHORS De Buyl,E., Lahaye,A., Ledoux,P. and Detroz,R.  
TITLE Xylanase, microorganisms producing it, DNA molecules, methods for preparing this xylanase and uses of the latter  
JOURNAL Patent: US 6346407-A 4 12-FEB-2002;  
FEATURES Location/Qualifiers  
source  
1..744  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 49.2%; Score 744; DB 6; Length 744;  
Best Local Similarity 100.0%; Pred. No. 3.9e-140; Mismatches 0; Indels 0; Gaps 0;  
Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 620 ATGAGACAAAAGAAATGACGTTGATTTTAGCCTTTTATAGCTTTTATAGTTTGTGCACTAACCTTA 679  
Db 1 ATGAGACAAAAGAAATGACGTTGATTTTAGCCTTTTATAGCTTTTATAGTTTGTGCACTAACCTTA 60  
Qy 680 CCTGAGAAATAATTCAGGCACAAATCGTCACCGCAATTCATTCGGCAACACGATGCG 739  
Db 61 CCTGAGAAATAATTCAGGCACAAATCGTCACCGCAATTCATTCGGCAACACGATGCG 120  
Qy 740 TATGATTATGAATTTTGGAAAGATAGCGGTGGCTCTCGGACAATGATTCTCAATCATGCG 799



Db 121 TATGATTATGAATTTTGGAAAGATAGCGTGGCTCTGGGACAAATGATTTCTCAATCATGCG 180  
Qy 800 GGTAGCTTCAGTCCCAATGGAACAAATGTTAAACATATTTATTCGTTAAAGGTAAAAA 859  
Db 181 GGTAGCTTCAGTCCCAATGGAACAAATGTTAAACATATTTATTCGTTAAAGGTAAAAA 240  
Qy 860 TTCAATGAACACAAACACCAACCAAGTTGGTAAACATGTCATAACTACGAGCCAAAC 919  
Db 241 TTCAATGAACACAAACACCAACCAAGTTGGTAAACATGTCATAACTACGAGCCAAAC 300  
Qy 920 TTCCAAACCAATGTTAAATGCGTATTTATGCGTCTATGTTGGACTGTTGACCCCTCTTGTG 979  
Db 301 TTCCAAACCAATGTTAAATGCGTATTTATGCGTCTATGTTGGACTGTTGACCCCTCTTGTG 360  
Qy 980 GAATATTTATTTGTCGACAGTTGGGCAACTGGCGTCCACAGGAGCAACGGCTTAAGGGG 1039  
Db 361 GAATATTTATTTGTCGACAGTTGGGCAACTGGCGTCCACAGGAGCAACGGCTTAAGGGG 420  
Qy 1040 ACCATCACTGTTGATGGAGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 1099  
Db 421 ACCATCACTGTTGATGGAGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480  
Qy 1100 TCCATTAAGGGGATTCGCCACATTTAAACAAATTTGGAGTGTTCGAAAGATCGAAACGCAAC 1159  
Db 481 TCCATTAAGGGGATTCGCCACATTTAAACAAATTTGGAGTGTTCGAAAGATCGAAACGCAAC 540  
Qy 1160 AGTGCCACGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAAACCTTAGGGATGAATATG 1219  
Db 541 AGTGCCACGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAAACCTTAGGGATGAATATG 600  
Qy 1220 GGGAAAATGTGAAGTCCGCTTACTGTAGAGGCTTCAAGTACGCGGAGTGCCTAAT 1279  
Db 601 GGGAAAATGTGAAGTCCGCTTACTGTAGAGGCTTCAAGTACGCGGAGTGCCTAAT 660  
Qy 1280 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTATGACGAG 1339  
Db 661 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTATGACGAG 720  
Qy 1340 AGCATAACTTTGGATAAAAAACAAT 1363  
Db 721 AGCATAACTTTGGATAAAAAACAAT 744

RESULT 8  
LOCUS AR193052 744 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 5 from patent US 6346407.  
ACCESSION AR193052  
VERSION AR193052.1 GI:20239017  
KEYWORDS Unknown.  
SOURCE Unclassified.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 744)  
AUTHORS De Buyl, E., Lahaye, A., Ledoux, P. and Detroz, R.  
TITLE Xylanase, microorganisms producing it, DNA molecules, methods for preparing this xylanase and uses of the latter  
JOURNAL Patent: US 6346407-A 5 12-FEB-2002;  
FEATURES Location/Qualifiers  
source 1..744  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 49.28; Score 744; DB 6; Length 744;  
Best Local Similarity 100.0%; Pred. No. 3.9e-140;  
Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 620 ATGACAAAGAAATGACGTTGATTTTAGCCCTTTTATGTTTGTGCACTAACCTTA 679  
Db 1 ATGACAAAGAAATGACGTTGATTTTAGCCCTTTTATGTTTGTGCACTAACCTTA 60

Qy 680 CCTGAGAAATTAATTCAGGCAAAATCGTCAACGCAAAATTCATTTGGCAACACCAATGCG 739  
Db 61 CCTGAGAAATTAATTCAGGCAAAATCGTCAACGCAAAATTCATTTGGCAACACCAATGCG 120  
Qy 740 TATGATTATGAATTTTGGAAAGATAGCGTGGCTCTGGGACAAATGATTTCTCAATCATGCG 799  
Db 121 TATGATTATGAATTTTGGAAAGATAGCGTGGCTCTGGGACAAATGATTTCTCAATCATGCG 180  
Qy 800 GGTAGCTTCAGTCCCAATGGAACAAATGTTAAACATATTTATTCGTTAAAGGTAAAAA 859  
Db 181 GGTAGCTTCAGTCCCAATGGAACAAATGTTAAACATATTTATTCGTTAAAGGTAAAAA 240  
Qy 860 TTCAATGAACACAAACACCAACCAAGTTGGTAAACATGTCATAACTACGAGCCAAAC 919  
Db 241 TTCAATGAACACAAACACCAACCAAGTTGGTAAACATGTCATAACTACGAGCCAAAC 300  
Qy 920 TTCCAAACCAATGTTAAATGCGTATTTATGCGTCTATGTTGGACTGTTGACCCCTCTTGTG 979  
Db 301 TTCCAAACCAATGTTAAATGCGTATTTATGCGTCTATGTTGGACTGTTGACCCCTCTTGTG 360  
Qy 980 GAATATTTATTTGTCGACAGTTGGGCAACTGGCGTCCACAGGAGCAACGGCTTAAGGGG 1039  
Db 361 GAATATTTATTTGTCGACAGTTGGGCAACTGGCGTCCACAGGAGCAACGGCTTAAGGGG 420  
Qy 1040 ACCATCACTGTTGATGGAGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 1099  
Db 421 ACCATCACTGTTGATGGAGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480  
Qy 1100 TCCATTAAGGGGATTCGCCACATTTAAACAAATTTGGAGTGTTCGAAAGATCGAAACGCAAC 1159  
Db 481 TCCATTAAGGGGATTCGCCACATTTAAACAAATTTGGAGTGTTCGAAAGATCGAAACGCAAC 540  
Qy 1160 AGTGCCACGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAAACCTTAGGGATGAATATG 1219  
Db 541 AGTGCCACGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAAACCTTAGGGATGAATATG 600  
Qy 1220 GGGAAAATGTGAAGTCCGCTTACTGTAGAGGCTTCAAGTACGCGGAGTGCCTAAT 1279  
Db 601 GGGAAAATGTGAAGTCCGCTTACTGTAGAGGCTTCAAGTACGCGGAGTGCCTAAT 660  
Qy 1280 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTATGACGAG 1339  
Db 661 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTATGACGAG 720  
Qy 1340 AGCATAACTTTGGATAAAAAACAAT 1363  
Db 721 AGCATAACTTTGGATAAAAAACAAT 744

RESULT 9  
LOCUS A68006 744 bp DNA linear PAT 05-MAY-1999  
DEFINITION Sequence 1 from Patent WO9743409.  
ACCESSION A68006  
VERSION A68006.1 GI:4756810  
KEYWORDS unidentified  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 744)  
AUTHORS Dalboe, H., Diderichsen, B., Sandal, T. and Kauppinen, S.  
TITLE METHOD OF PROVIDING NOVEL DNA SEQUENCES  
JOURNAL Patent: WO 9743409-A 1 20-NOV-1997;  
NOVONORDISK AS (DK)  
FEATURES Location/Qualifiers  
source 1..744  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/strain="BACILLUS SP. AC13, NCIMB NO. 40482"  
/db\_xref="taxon:32644"  
1..>744  
/notes="unnamed protein product"  
/codon\_start=1  
CDS

/protein\_id="CAB42305.1"  
/db\_xref="GI:4756811"  
/translation="WRKKLFLFLAFLVCFALTLPALIIQAIQIVTDNSIGNHGDGVDE  
FWKDSGGSTWLNHGGTFSQWNNNLLFRKKKFNETHQVQGNMNSINYANFQ  
SNGAYLCVYGNWDPVLEYIYVDSWGNRPFGATPKGTITVDGTYDIETLRNQP  
SNGIATFKQWSVRSRKSTGTTISVSNFRANWLNMGWMEVALTVBEGYSSGS  
ANVYSNLRINGNPLSTISNDKSITLDKNN"

ORIGIN	Query Match	47.2%; Score 713.6; DB 6; Length 744;
	Best Local Similarity 97.4%; Pred. No. 5.4e-134;	
	Matches 725; Conservative 0; Mismatches 19; Indels 0; Gaps 0;	
QY	620	ATGAGACAAAAGAAATTCAGCTTGATTTTAGCTTTTGGCTTTTGGCTTGTGCTACTAACCTTA 679
DB	1	ATGAGACAAAAGAAATTCAGCTTGATTTTAGCTTTTGGCTTTTGGCTTGTGCTACTAACCTTA 60
QY	680	CCTGCAGAAATAATTCAGGCACAAATCGTCACCGCAATTCATTTGGCAACCCAGCATGC 739
DB	61	CCTGCAGAAATAATTCAGGCACAAATCGTCACCGCAATTCATTTGGCAACCCAGCATGC 120
QY	740	TATGATTATGAATTTTGGAAAGATAGCGTGGCTCTGGGACAATGATTTCTCAATCATGC 799
DB	121	TATGATTATGAATTTTGGAAAGATAGCGTGGCTCTGGGACAATGATTTCTCAATCATGC 180
QY	800	GGTACGTTTCAGTCCCAATGGAACAATGTTAACACATATTTCCGTAAAGGTAAAAA 859
DB	181	GTTACGTTTCAGTCCCAATGGAACAATGTTAACACATATTTCCGTAAAGGTAAAAA 240
QY	860	TTCAATGAACAACAACAACACCAACAGTTGGTAAACATGTCATAAACTACGGAGCCAA 919
DB	241	TTCAATGAACAACAACAACACCAACAGTTGGTAAACATGTCATAAACTACGGAGCCAA 300
QY	920	TTCCAAACCAATGTTAATGCGTATTTATGCGTCTATGTTGACCTCTTGTG 979
DB	301	TTCCAGCCAAACGGAAATGCGTATTTATGCGTCTATGTTGACCTCTTGTG 360
QY	980	GAATATTATATTGTCGACAGTTGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGG 1039
DB	361	GAATATTATATTGTCGACAGTTGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGG 420
QY	1040	ACCATCACTGTTGATGGAGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACC 1099
DB	421	ACCATCACTGTTGATGGAGAACATATGATATCTATGAACCTCTTAGAGTCAATCAGCC 480
QY	1100	TCCATTAAGGGGATTCACATTTAAACAATTTGGAGTGTTCGAAGATCGAAACGACG 1159
DB	481	TCCATTAAGGGGATTCACATTTAAACAATTTGGAGTGTTCGAAGATCGAAACGACG 540
QY	1160	AGTGGCAGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAAACCTTAGGGATGAATG 1219
DB	541	AGTGGCAGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAAACCTTAGGGATGAATG 600
QY	1220	GGGAAAATGATGAAGTCGCGCTTACTGTAGAAGGCTATCAAGTAGCGGAAGTCTAAT 1279
DB	601	GGGAAAATGATGAAGTCGCGCTTACTGTAGAAGGCTATCAAGTAGCGGAAGTCTAAT 660
QY	1280	GTATATAGCAATACACTAAGTAACCGTAAACCTCTCTCACTATTAGTAATGACGAG 1339
DB	661	GTATATAGCAATACACTAAGTAACCGTAAACCTCTCTCACTATTAGTAATGACGAG 720
QY	1340	AGCAATACTTTGGATAAAAAACAAT 1363
DB	721	AGCAATACTTAGATAAAAAACAAT 744

RESULT 10  
LOCUS AR163110 744 bp DNA linear PAT 17-OCT-2001  
DEFINITION Sequence 1 from patent US 6270968.  
ACCESSION AR163110  
VERSION AR163110.1 GI:16233600  
KEYWORDS

SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	ORIGIN	Query Match	47.2%; Score 713.6; DB 6; Length 744;
Unknown.	Unknown.	Unclassified	Unclassified	1 (bases 1 to 744)	Patent: US 6270968-A 1 07-AUG-2001;	Location/Qualifiers		Best Local Similarity 97.4%; Pred. No. 5.4e-134;	
					interest	1. 744		Matches 725; Conservative 0; Mismatches 19; Indels 0; Gaps 0;	
						/organism="unknown"			
						/mol_type="unassigned DNA"			
QY							620	ATGAGACAAAAGAAATTCAGCTTGATTTTAGCTTTTGGCTTTTGGCTTGTGCTACTAACCTTA 679	
DB							1	ATGAGACAAAAGAAATTCAGCTTGATTTTAGCTTTTGGCTTTTGGCTTGTGCTACTAACCTTA 60	
QY							680	CCTGCAGAAATAATTCAGGCACAAATCGTCACCGCAATTCATTTGGCAACCCAGCATGC 739	
DB							61	CCTGCAGAAATAATTCAGGCACAAATCGTCACCGCAATTCATTTGGCAACCCAGCATGC 120	
QY							740	TATGATTATGAATTTTGGAAAGATAGCGTGGCTCTGGGACAATGATTTCTCAATCATGC 799	
DB							121	TATGATTATGAATTTTGGAAAGATAGCGTGGCTCTGGGACAATGATTTCTCAATCATGC 180	
QY							800	GGTACGTTTCAGTCCCAATGGAACAATGTTAACACATATTTCCGTAAAGGTAAAAA 859	
DB							181	GTTACGTTTCAGTCCCAATGGAACAATGTTAACACATATTTCCGTAAAGGTAAAAA 240	
QY							860	TTCAATGAACAACAACAACACCAACAGTTGGTAAACATGTCATAAACTACGGAGCCAA 919	
DB							241	TTCAATGAACAACAACAACACCAACAGTTGGTAAACATGTCATAAACTACGGAGCCAA 300	
QY							920	TTCCAAACCAATGTTAATGCGTATTTATGCGTCTATGTTGACCTCTTGTG 979	
DB							301	TTCCAGCCAAACGGAAATGCGTATTTATGCGTCTATGTTGACCTCTTGTG 360	
QY							980	GAATATTATATTGTCGACAGTTGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGG 1039	
DB							361	GAATATTATATTGTCGACAGTTGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGG 420	
QY							1040	ACCATCACTGTTGATGGAGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACC 1099	
DB							421	ACCATCACTGTTGATGGAGAACATATGATATCTATGAACCTCTTAGAGTCAATCAGCC 480	
QY							1100	TCCATTAAGGGGATTCACATTTAAACAATTTGGAGTGTTCGAAGATCGAAACGACG 1159	
DB							481	TCCATTAAGGGGATTCACATTTAAACAATTTGGAGTGTTCGAAGATCGAAACGACG 540	
QY							1160	AGTGGCAGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAAACCTTAGGGATGAATG 1219	
DB							541	AGTGGCAGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAAACCTTAGGGATGAATG 600	
QY							1220	GGGAAAATGATGAAGTCGCGCTTACTGTAGAAGGCTATCAAGTAGCGGAAGTCTAAT 1279	
DB							601	GGGAAAATGATGAAGTCGCGCTTACTGTAGAAGGCTATCAAGTAGCGGAAGTCTAAT 660	
QY							1280	GTATATAGCAATACACTAAGTAACCGTAAACCTCTCTCACTATTAGTAATGACGAG 1339	
DB							661	GTATATAGCAATACACTAAGTAACCGTAAACCTCTCTCACTATTAGTAATGACGAG 720	
QY							1340	AGCAATACTTTGGATAAAAAACAAT 1363	
DB							721	AGCAATACTTAGATAAAAAACAAT 744	

RESULT 11

A45313	A45313		744 bp	DNA	linear	PAT 07-MAR-1997
LOCUS	Sequence 18 from Patent WO9518219.					
DEFINITION	A45313					
ACCESSION	A45313					
VERSION	A45313.1	GI:2299796				
KEYWORDS	unidentified					
SOURCE	unidentified					
ORGANISM	unclassified					
REFERENCE	1 (bases 1 to 744)					
AUTHORS	Van,S.P., Williams,D.P., Iverson,S., Farrell,R.L., Herbes,W.T., Van,D.K., Herweijer,M.A., Van,B.R., Quax,W.J., Goedegebuur,P. and Jones,B.E.					
TITLE	ALKALI-TOLERANT XYLANASES					
JOURNAL	Patent: WO 9518219-A 18 06-JUL-1995; GIST BROCADES NV (NL)					
COMMENT	Other publication JP 8507221T 960806 Other publication BR 9405934 951226 Other publication NO 95312 951019 Other publication FI 953920 950821 Other publication AU 1415095 950717. Location/Qualifiers					
FEATURES	1..744					
source	/organism="unidentified"					
	/mol_type="unassigned DNA"					
	/strain="I-43-3"					
	/isolate="CBS672.93"					
	/db_xref="taxon:32644"					
CDS	1..744					
	/codon_start=1					
	/product="xyLANASE"					
	/protein_id="CAA02821.1"					
	/db_xref="GI:2299797"					
	/translation="MSOKKLTLLNLSFALTLPARISQAIIVTDNSIATRGGYDYEF WKDGGSGWTMLNHGGTFSAQMNNVNIIPRKGNKTETHOQVGNSINYGANQP NGNYLCYCVGTVDPLVEYYIVDSKGWNPFGATPKGTITVDSGTYDIIVTLRVNPS IKGAIPTKWVSRRSKTSGETISVNHFRAWENLGMNMKNYEVALTVEGYQSSSSA NVYSNTLRINGNLSTISSNESITLDKNN"					
ORIGIN						
Query Match	44.5%	Score 673.6;	DB 6;	Length 744;		
Best Local Similarity	95.7%;	Pred. No. 6.5e-126;				
Matches 714;	Conservative 0;	Mismatches 29;	Indels 3;	Gaps 2;		
Qy	620	ATGCACAAAGAATAATTGAGCTTGTAATTTTAGCCCTTTTTAGTTGTTGTTTGCATTAACCTTA	679			
Db	1	ATGAGCCAAAGAATAATTGACGTTCGA--TTAACCTTTTTAGTTTG--TTTGCACTTAACCTTA	57			
Qy	680	CTTGCAGAAATAATTACAGGCCACAATACTCGTCAACGACAAATTCATTCAGCACACCAGCATGCC	739			
Db	58	CCTGCAAGAATAAAGTCAGGCCACAATACTCGTCAACGACAAATTCATTCAGCACACCAGCGGTGGT	117			
Qy	740	TATGATTTATGATTTTTCGARAAGTAGCGGTGGCTCTCGGACAAATGATTTCTCAATCATGGC	799			
Db	118	TATGATTTATGATTTTTCGAAAAGATAGCGGTGGCTCTCGGACAAATGATTTCTCAATCATGGC	177			
Qy	800	GGTACGTTTCAGTGCCTCCCAATGGAAACAATGTTTAAACAACATATATTATTCCTGTAAGGTAAAAA	859			
Db	178	GGTACGTTTCAGTGCCTCCCAATGGAAACAATGTTTAAACAACATATATTATTCCTGTAAGGTAAAAA	237			
Qy	860	TTCAATGAAACACAAACACACCAACAAGTGGTAAATGTCATTAACATTAACATCGGAGCCAAC	919			
Db	238	TTCAATGAAACACAAACACACCAACAAGTGGTAAATGTCATTAACATTAACATCGGAGCCAAC	297			
Qy	920	TTCAACCAATGTTAATGCGTATTTATGCGTCTATGGTTGGACTGTTGACCCCTCTTGTC	979			
Db	298	TTCAAGCCAAACGGTAAATGCGTATTTATGCGTCTATGGTTGGACTGTTGACCCCTCTTGTC	357			
Qy	980	GAATATTATATTGTCGACAGTTGGGGCAACTGGCGTCCACGAGGACAAACGCTTAAGGGG	1039			
Db	358	GAATATTATATTGTCGACAGTTGGGGCAACTGGCGTCCACGAGGACAAACGCTTAAGGGG	417			
Qy	1040	ACCATCACTGTTGATGGAGGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC	1099			

Db 298 TTCCAGCCAAACGGTAATCGTATTTATCGCTATGTTGAGCTGTTGACCCCTCTGTT 357  
 QY 980 GAATATTATTTGTCGACAGTTGGGCACTGGCGTCCACAGGAGCAACGCTTAAGGG 1039  
 Db 358 GAAATATTATTTGTCGACAGTTGGGCACTGGCGTCCACAGGAGCAACGCTTAAGGGA 417  
 QY 1040 ACCATCACTGTTGATGGAGGAACATATGATATCTACGAGACTCTTTAGAGTCAATCAAGCC 1099  
 Db 418 ACCATCACTGTTGATGGAGGAACATATGATATCTAGAACTCTTAGAGTCAATCAGCC 477  
 QY 1100 TCATTAAGGGATTCGCACATTTAAACATATTCGAGTGTTCGAAGATCGAAACGACG 1159  
 Db 478 TCCATTAAGGGATTCGCACATTTAAACATATTCGAGTGTTCGAAGATCGAAACGACG 537  
 QY 1160 AGTGGCAGCATTTCTGTAGCAACCACTTTAGAGCTGGGAAACTTTAGGATGAATATG 1219  
 Db 538 AGTGGCAGCATTTCTGTAGCAACCACTTTAGAGCTGGGAAACTTTAGGATGAATATG 597  
 QY 1220 GGGAAATGATGAAGTGCCTTACTGTAGAGGCTATCAAGTAGCGGAAGTCTTAAT 1279  
 Db 598 GGGAAATGATGAAGTGCCTTACTGTAGAGGCTATCAAGTAGCGGAAGTCTTAAT 657  
 QY 1280 GTATATAGCAATACACTAAGATTAACGGTAACCTCTCTCACTATTAGTAATGACGAG 1339  
 Db 658 GTATATAGCAATACACTAAGATTAACGGTAACCTCTCTCACTATTAGTAATGACGAG 717  
 QY 1340 AGCATAACTTTGGATAAAACAAATTA 1365  
 Db 718 AGCATAACTCTAGATAAAACAAATTA 743

RESULT 13  
 A48222  
 LOCUS A48222  
 DEFINITION Sequence 1 from Patent EP0698667.  
 ACCESSION A48222  
 VERSION A48222.1 GI:2302069  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 unclassified  
 unclassified  
 unclassified

REFERENCE 1 (bases 1 to 663)  
 De,B.E., Lahaye,A., Ledoux,P. and Detroz,R.  
 Xylanase, microorganisms for its production, DNA molecules, process  
 of preparation and use thereof  
 Patent: EP 0698667-A 1 28-FEB-1996;  
 JOURNAL SOLVAY (BE)  
 COMMENT Other publication BE 1008751 960702  
 Other publication BE 1008570 960604  
 Other publication BR 9503454 960305  
 Other publication JP 8092284 960409  
 Other publication FI 953578 960127  
 Other publication CA 2154628 960127  
 Other publication AU 2508695 960208.

FEATURES  
 source  
 1..663  
 /organism="unclassified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"

ORIGIN  
 Query Match 43.8%; Score 663; DB 6; Length 663;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-124;  
 Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 701 CAAATCGTCACGCAATTCATTCGCAACCAAGTGGCTATGATTGAATTTGGAA 760  
 Db 1 CAAATCGTCACGCAATTCATTCGCAACCAAGTGGCTATGATTGAATTTGGAA 60  
 QY 761 GATACGGTGGCTCTGGGACATGATTTCTCAATCATGGCGGTACGTTCAAGTGCACATGG 820  
 Db 61 GATACGGTGGCTCTGGGACATGATTTCTCAATCATGGCGGTACGTTCAAGTGCACATGG 120

QY 821 AACCAATGTTAAACAATATTATTCGTAAGGTAAAAAATTCATATGAACAACAACACAC 880  
 Db 121 AACCAATGTTAAACAATATTATTCGTAAGGTAAAAAATTCATATGAACAACAACACAC 180  
 QY 881 CAAACAAGTTGGTAAACATGTCATATAAATACGAGGCCAACTTCCAAACCAATATGGTAATGCG 940  
 Db 181 CAAACAAGTTGGTAAACATGTCATATAAATACGAGGCCAACTTCCAAACCAATATGGTAATGCG 240  
 QY 941 TATTATGCGTCTATGTTGGACTGTTGACCTCTCTCGAATATTTATTTGTCGACAGT 1000  
 Db 241 TATTATGCGTCTATGTTGGACTGTTGACCTCTCTCGAATATTTATTTGTCGACAGT 300  
 QY 1001 TGGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 1060  
 Db 301 TGGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 360  
 QY 1061 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTTGCCACA 1120  
 Db 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTTGCCACA 420  
 QY 1121 TTTAAACAATATTCGAGTGTTCGAAGATCGAAACGCGACGAGTGGCAGATTTCTGTGTCAGC 1180  
 Db 421 TTTAAACAATATTCGAGTGTTCGAAGATCGAAACGCGACGAGTGGCAGATTTCTGTGTCAGC 480  
 QY 1181 AACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATGGGAAATGTATGAAGTGGCG 1240  
 Db 481 AACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATGGGAAATGTATGAAGTGGCG 540  
 QY 1241 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTGGCTAATGTATATAGCAATACACTAAGA 1300  
 Db 541 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTGGCTAATGTATATAGCAATACACTAAGA 600  
 QY 1301 ATTAACGTTAAACCTCTCTCAACTATTAGTAATGACGAGAGCATAACTTTGGATAAAAC 1360  
 Db 601 ATTAACGTTAAACCTCTCTCAACTATTAGTAATGACGAGAGCATAACTTTGGATAAAAC 660  
 QY 1361 AAT 1363  
 Db 661 AAT 663

RESULT 14  
 A48223  
 LOCUS A48223  
 DEFINITION Sequence 2 from Patent EP0698667.  
 ACCESSION A48223  
 VERSION A48223.1 GI:2302070  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 unclassified  
 unclassified  
 unclassified

REFERENCE 1 (bases 1 to 663)  
 De,B.E., Lahaye,A., Ledoux,P. and Detroz,R.  
 Xylanase, microorganisms for its production, DNA molecules, process  
 of preparation and use thereof  
 Patent: EP 0698667-A 2 28-FEB-1996;  
 JOURNAL SOLVAY (BE)  
 COMMENT Other publication BE 1008751 960702  
 Other publication BE 1008570 960604  
 Other publication BR 9503454 960305  
 Other publication JP 8092284 960409  
 Other publication FI 953578 960127  
 Other publication CA 2154628 960127  
 Other publication AU 2508695 960208.

FEATURES  
 source  
 1..663  
 /organism="unclassified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"

CDS  
 1..663  
 /notes="unnamed protein product; Protein sequence is in  
 conflict with the conceptual translation"  
 /codon\_start=1

/protein\_id="CAA03091.1"  
/db\_xref="GI:2302071"  
/translation="QIVTDSIGNHDGVDYFWDKSGSGTMIILNHGCTFPAQNNVN  
NIFRKKFKNETQHQVGNMISYINFGPNMAYLCVYGVTDVPEVYIIVDSWG  
NWRPGATPKGTITVDGGTIDYETLRVNFQSIKGIATFKQYWSVRKRTSGTISV  
NHFRWENLGMNMGRWYEVALTVEGYQSGSANVYNTLRINGNPLSTISNDESIITLD  
KNN"  
mat\_peptide 1..663  
/product="unnamed"

ORIGIN		Query Match	43.8%; Score 663; DB 6; Length 663;
		Best Local Similarity	100.0%; Pred. No. 9.1e-124;
		Matches 663; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	701	CAAAATCGTCACCGACAAATTCATTTGGCAACACGATGGCTATGATATGAATTTTGGAAA	760
Db	1	CAAAATCGTCACCGACAAATTCATTTGGCAACACGATGGCTATGATATGAATTTTGGAAA	60
QY	761	GATAGCGGTGGCTCTGGGACAATGATCTCAATCATGCGGCTACGTTCAAGTCCCAATGG	820
Db	61	GATAGCGGTGGCTCTGGGACAATGATCTCAATCATGCGGCTACGTTCAAGTCCCAATGG	120
QY	821	AACAATGTTAAACAACATATTTATTCGTTAAAGGTAAAAAATTCATGAAAAACACACAC	880
Db	121	AACAATGTTAAACAACATATTTATTCGTTAAAGGTAAAAAATTCATGAAAAACACACAC	180
QY	881	CAACAAGTTGGTAACATGTCATAAATACTACGAGCCAACTTCCAAACAAATGGTAATGG	940
Db	181	CAACAAGTTGGTAACATGTCATAAATACTACGAGCCAACTTCCAAACAAATGGTAATGG	240
QY	941	TATTTATGGCTCTATGGTTGGACTGTGACCCCTCTTGTGCAATATTTATTTGCGACAGT	1000
Db	241	TATTTATGGCTCTATGGTTGGACTGTGACCCCTCTTGTGCAATATTTATTTGCGACAGT	300
QY	1001	TGGGGCAACTGGCGTCCACGAGCAACGCCCTTAAGGGGACCATCACTGTGATGGAGGA	1060
Db	301	TGGGGCAACTGGCGTCCACGAGCAACGCCCTTAAGGGGACCATCACTGTGATGGAGGA	360
QY	1061	ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATGGCCACA	1120
Db	361	ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATGGCCACA	420
QY	1121	TTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGGATTTCTGTGAGC	1180
Db	421	TTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGGATTTCTGTGAGC	480
QY	1181	AACCACTTTAGAGCGTGGGAAAACCTTAGGATGATATGCGGAAAATGTATGAAGTCGCG	1240
Db	481	AACCACTTTAGAGCGTGGGAAAACCTTAGGATGATATGCGGAAAATGTATGAAGTCGCG	540
QY	1241	CTTACTGTAGAAGCTATCAAGATAGCGGAAGTGTCTAATGATATAGCAATACACTAAGA	1300
Db	541	CTTACTGTAGAAGCTATCAAGATAGCGGAAGTGTCTAATGATATAGCAATACACTAAGA	600
QY	1301	ATTAAACGGTAACCCCTCTCAACTATTTAGTAATGACGAGACATACCTTTGGATAAAAC	1360
Db	601	ATTAAACGGTAACCCCTCTCAACTATTTAGTAATGACGAGACATACCTTTGGATAAAAC	660
QY	1361	AAT 1363	
Db	661	AAT 663	

RESULT 15  
AR193049 LOCUS AR193049 563 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 1 from patent US 6346407.  
ACCESSION AR193049  
VERSION AR193049.1 GI:20239014  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

Unclassified.  
1 (bases 1 to 663)  
De Buyl, E., Lahaye, A., Ledoux, P. and Detroz, R.  
Xylanase, microorganisms producing it, DNA molecules, methods for  
preparing this xylanase and uses of the latter  
Patent: US 6346407-A 1 12-FEB-2002;  
Location/Qualifiers  
1..663  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN		Query Match	43.8%; Score 663; DB 6; Length 663;
		Best Local Similarity	100.0%; Pred. No. 9.1e-124;
		Matches 663; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	701	CAAAATCGTCACCGACAAATTCATTTGGCAACACGATGGCTATGATATGAATTTTGGAAA	760
Db	1	CAAAATCGTCACCGACAAATTCATTTGGCAACACGATGGCTATGATATGAATTTTGGAAA	60
QY	761	GATAGCGGTGGCTCTGGGACAATGATCTCAATCATGCGGCTACGTTCAAGTCCCAATGG	820
Db	61	GATAGCGGTGGCTCTGGGACAATGATCTCAATCATGCGGCTACGTTCAAGTCCCAATGG	120
QY	821	AACAATGTTAAACAACATATTTATTCGTTAAAGGTAAAAAATTCATGAAAAACACACAC	880
Db	121	AACAATGTTAAACAACATATTTATTCGTTAAAGGTAAAAAATTCATGAAAAACACACAC	180
QY	881	CAACAAGTTGGTAACATGTCATAAATACTACGAGCCAACTTCCAAACAAATGGTAATGG	940
Db	181	CAACAAGTTGGTAACATGTCATAAATACTACGAGCCAACTTCCAAACAAATGGTAATGG	240
QY	941	TATTTATGGCTCTATGGTTGGACTGTGACCCCTCTTGTGCAATATTTATTTGCGACAGT	1000
Db	241	TATTTATGGCTCTATGGTTGGACTGTGACCCCTCTTGTGCAATATTTATTTGCGACAGT	300
QY	1001	TGGGGCAACTGGCGTCCACGAGCAACGCCCTTAAGGGGACCATCACTGTGATGGAGGA	1060
Db	301	TGGGGCAACTGGCGTCCACGAGCAACGCCCTTAAGGGGACCATCACTGTGATGGAGGA	360
QY	1061	ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATGGCCACA	1120
Db	361	ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATGGCCACA	420
QY	1121	TTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGGATTTCTGTGAGC	1180
Db	421	TTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGGATTTCTGTGAGC	480
QY	1181	AACCACTTTAGAGCGTGGGAAAACCTTAGGATGATATGCGGAAAATGTATGAAGTCGCG	1240
Db	481	AACCACTTTAGAGCGTGGGAAAACCTTAGGATGATATGCGGAAAATGTATGAAGTCGCG	540
QY	1241	CTTACTGTAGAAGCTATCAAGATAGCGGAAGTGTCTAATGATATAGCAATACACTAAGA	1300
Db	541	CTTACTGTAGAAGCTATCAAGATAGCGGAAGTGTCTAATGATATAGCAATACACTAAGA	600
QY	1301	ATTAAACGGTAACCCCTCTCAACTATTTAGTAATGACGAGACATACCTTTGGATAAAAC	1360
Db	601	ATTAAACGGTAACCCCTCTCAACTATTTAGTAATGACGAGACATACCTTTGGATAAAAC	660
QY	1361	AAT 1363	
Db	661	AAT 663	

Search completed: November 10, 2004, 21:26:31  
Job time : 7322.01 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 10, 2004, 15:55:17 ; Search time 959.753 Seconds  
(without alignments)  
8275.436 Million cell updates/sec

Title: US-09-909-207-10  
Perfect score: 1513  
Sequence: 1 AAATTGAATTGGTATATCT.....TGAAACACCTCGTCACTAG 1513

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4134886 seqs, 2624710521 residues  
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 23Sep04: \*  
1: Geneseqn1980s: \*  
2: Geneseqn1990s: \*  
3: Geneseqn2000s: \*  
4: Geneseqn2001as: \*  
5: Geneseqn2001bs: \*  
6: Geneseqn2002as: \*  
7: Geneseqn2002bs: \*  
8: Geneseqn2003as: \*  
9: Geneseqn2003bs: \*  
10: Geneseqn2003cs: \*  
11: Geneseqn2003ds: \*  
12: Geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1513	100.0	1513	2	AAT16103
2	840.6	55.6	871	2	AAT16103 Xylanase
3	744	49.2	744	2	AAT16102
4	713.6	47.2	744	2	AAT16102 Xylanase
5	673.6	44.5	744	2	AAT16102 Glycosyl
6	663	43.8	663	2	AAT16101
7	590.2	39.0	747	2	AAT16101 Xylanase
8	393	26.0	1068	12	AAT13074 Xylanase
9	282.8	18.7	1022	2	Adj34947 DNA encod
10	276.6	18.3	1956	12	AAT09023 B. pumilu
11	274.4	18.1	684	12	Adj35051 DNA encod
12	255.6	16.9	2364	3	Adj34969 DNA encod
13	217.2	14.4	1190	2	Aaz51821 Clostridi
14	204.8	13.5	1244	2	Aat90972 Nucleotid
15	177.8	11.8	747	12	Aat08142 Xylanase
16	154.4	10.2	164	2	Adj35011 DNA encod
17	151.2	10.0	164	2	Aaq92876 Thermosta
18	143.6	9.5	1695	12	Aaq92875 Thermosta
19	142.6	9.4	1338	12	Adj35101 DNA encod
20	136.2	9.0	1077	12	Adj35151 DNA encod
21	130.2	8.6	1065	12	Adj34965 DNA encod

22	128.4	8.5	1047	12	ADJ34949
23	121.2	8.0	678	12	ADJ34955
24	119.4	7.9	1375	2	AAx90405
25	119.4	7.9	1375	2	AAx90405 Actinomad
26	119	7.9	229	2	AAv15063 Xylanase
27	118.4	7.8	573	2	AAv36098 DNA seque
28	117.8	7.8	1207	2	AAv36098 DNA seque
29	117.6	7.8	1041	12	AAv36098 DNA seque
30	116.8	7.7	1047	12	AAv36098 DNA seque
31	115.2	7.6	213	2	AAv15059 Xylanase
32	115.2	7.6	596	3	AAv15059 Xylanase
33	115.2	7.6	596	6	AAv15059 Xylanase
34	115.2	7.6	596	9	AAv15059 Xylanase
35	115.2	7.6	596	9	AAv15059 Xylanase
36	112	7.4	1074	12	AAv15059 Xylanase
37	112	7.4	1137	12	AAv15059 Xylanase
38	111.8	7.4	1273	2	AAv15059 Xylanase
39	111.4	7.4	636	12	AAv15059 Xylanase
40	111.4	7.4	942	10	AAv15059 Xylanase
41	111.2	7.3	1008	12	AAv15059 Xylanase
42	110.8	7.3	669	12	AAv15059 Xylanase
43	110.4	7.3	1047	12	AAv15059 Xylanase
44	110	7.3	1002	10	AAv15059 Xylanase
45	109.4	7.2	1041	12	AAv15059 Xylanase

## ALIGNMENTS

RESULT 1  
AAT16103  
ID AAT16103 standard; DNA; 1513 BP.  
XX AC AAT16103;  
XX 16-OCT-2003 (revised)  
DT 15-MAY-1996 (first entry)  
XX Xylanase gene.  
XX Xylanase; thermostable enzyme; paper; pulp; bleaching; feedstuff; baking;  
KW ss.  
XX Bacillus sp; strain 710/1 (LMG P-14798).  
XX Key Location/Qualifiers  
FT promoter 1..619  
FT CDS /\*tag= a  
FT /\*tag= b  
FT sig\_peptide 620..600  
FT /\*tag= c  
FT mat\_peptide 701..1363  
FT /\*tag= d  
FT 3'UTR /\*tag= e  
FT EC number= "3.2.1.8"  
FT 1367..1513  
FT /\*tag= e  
XX AU9525086-A.  
XX 08-FEB-1996.  
XX 19-JUL-1995; 95AU-00025086.  
XX 26-JUL-1994; 94BB-00000706.  
XX 17-MAY-1995; 95BB-00000448.  
XX (SOLV ) SOLVAY SA.  
XX De Buyl E, Lahaye A, Ledoux P, Detroz R;  
XX WPI; 1996-117341/13.  
XX P-PSDB; AAR92054.

XX Bacillus derived xylanase active over wide pH range - used in treatment  
PT of paper pulp, animal feeds and in bakery goods.  
XX Claim 10; Page 59-61; 94pp; English.  
XX A DNA sequence (AAT16102) coding for a xylanase precursor (AAR92054) was  
CC isolated from a gene library of *Bacillus* sp. 720/1 (IMG P-14798). The  
CC gene may be incorporated into a vector and expressed in transformed  
CC hosts, pref. *Bacillus licheniformis* or *Bacillus pumilus*, for prodn. of  
CC thermostable mature xylanase (AAR92053). The enzyme is useful in the  
CC paper pulp, animal feed and baking industries. (Updated on 16-OCT-2003 to  
CC standardise OS field)  
XX Sequence 1513 BP; 500 A; 282 C; 271 G; 460 T; 0 U; 0 Other;  
SQ Query Match 100.0%; Score 1513; DB 2; Length 1513;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATTGAATGTTGATATCTAATGATAGACAAATCGTCATCTGTTTAACTAATCTC 60  
DB 1. AAATTGAATGTTGATATCTAATGATAGACAAATCGTCATCTGTTTAACTAATCTC 60  
QY 61 AAACCAATCTCTTTTATTAAAGCTTAACCACTTCAATCTTATCACAAGAACATCTTT 120  
DB 61' AAACCAATCTCTTTTATTAAAGCTTAACCACTTCAATCTTATCACAAGAACATCTTT 120  
QY 121 APAGAACTTCCCAATTTGCAAGACGATAAAATCTTTTCCCTATTTTATCTTATCG 180  
DB 121 APAGAACTTCCCAATTTGCAAGACGATAAAATCTTTTCCCTATTTTATCTTATCG 180  
QY 181 CCGTGATCGTTTAAATTTGTAACCTTATTTAGTTAGTGATGTTCCCTCATCTAC 240  
DB 181 CCGTGATCGTTTAAATTTGTAACCTTATTTAGTTAGTGATGTTCCCTCATCTAC 240  
QY 241 CAATTAATCACAGTTAAGCTAGAGTCACTCTTTTTCGGTTCTCAAAATACCTGAAGAC 300  
DB 241 CAATTAATCACAGTTAAGCTAGAGTCACTCTTTTTCGGTTCTCAAAATACCTGAAGAC 300  
QY 301 ATTTATGTCATATTTCTCACGCCCTCCATAATGGAATATATATCTTTTATACATA 360  
DB 301 ATTTATGTCATATTTCTCACGCCCTCCATAATGGAATATATATCTTTTATACATA 360  
QY 361 TTAAGTAAATAGTATATCTTCAAGCTTCAATCAATGATGATGATGATGATGATGATGAT 420  
DB 361 TTAAGTAAATAGTATATCTTCAAGCTTCAATCAATGATGATGATGATGATGATGATGAT 420  
QY 421 AGCAGTATCCAAAAACACATGATGTTGACCTCTTTAAAGAAAGTGCATCTATGAAAA 480  
DB 421 AGCAGTATCCAAAAACACATGATGTTGACCTCTTTAAAGAAAGTGCATCTATGAAAA 480  
QY 481 GATAATATCCAGTTTCAAAATTTGAAATAGTGTGATGGAATAGTTTGAATGTCAACTG 540  
DB 481 GATAATATCCAGTTTCAAAATTTGAAATAGTGTGATGGAATAGTTTGAATGTCAACTG 540  
QY 541 CTGTGAAGAGGGTAGGTAGTACCGTAGACTTCAATACCAAAATAGTTGTTAAAAAA 600  
DB 541 CTGTGAAGAGGGTAGGTAGTACCGTAGACTTCAATACCAAAATAGTTGTTAAAAAA 600  
QY 601 TTTAAAGAGGAAATGCTAATCAGACAAAAGAAATTTGACGTTGATTTAGCCTTTTATG 660  
DB 601 TTTAAAGAGGAAATGCTAATCAGACAAAAGAAATTTGACGTTGATTTAGCCTTTTATG 660  
QY 661 TTGTTTTGCACTTAACCTTACCTGAGAAATAATTCAGGCACAAATCGTCACGCAATTC 720  
DB 661 TTGTTTTGCACTTAACCTTACCTGAGAAATAATTCAGGCACAAATCGTCACGCAATTC 720  
QY 721 CATTCGCAACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
DB 721 CATTCGCAACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
QY 781 AATGATTTCTCAATCATGCGGTACGTTTCAAGTCCCAATGGAACAATGTTAAACAATATT 840

DB 781 AATGATTTCTCAATCATGCGGTACGTTGAGTCCCAATGGAACAATGTTAAACAATATT 840  
QY 841 ATTCGGTAAAGGTAAAAATTTCAATGAAACACAAACACACAAAGTTGGTAAATGTC 900  
DB 841 ATTCGGTAAAGGTAAAAATTTCAATGAAACACAAACACACAAAGTTGGTAAATGTC 900  
QY 901 CATAACTACGAGCGCACTTCCAAACCAATGGAATGGAATGGAATGGAATGGAATGGAATG 960  
DB 901 CATAACTACGAGCGCACTTCCAAACCAATGGAATGGAATGGAATGGAATGGAATGGAATG 960  
QY 961 GACTGTTGACCTCTTGTGCAATATATATATGTCGACAGTTGGGCAACTGGCGTCCACC 1020  
DB 961 GACTGTTGACCTCTTGTGCAATATATATATGTCGACAGTTGGGCAACTGGCGTCCACC 1020  
QY 1021 AGGAGCAACGCTTAAGGGGACCATCTGTTGATGGAGGAAACATATGATCTACGAGAC 1080  
DB 1021 AGGAGCAACGCTTAAGGGGACCATCTGTTGATGGAGGAAACATATGATCTACGAGAC 1080  
QY 1081 TCTTAGAGTCAATCAACCTCCATTTAAGGGGATTCGCAATTTAAACAATATTGGAGTGT 1140  
DB 1081 TCTTAGAGTCAATCAACCTCCATTTAAGGGGATTCGCAATTTAAACAATATTGGAGTGT 1140  
QY 1141 TCGAAGATCGAAACGACGAGTGGCAGCATTTCTGTCAGCAACCACTTTAGAGCGTGGGA 1200  
DB 1141 TCGAAGATCGAAACGACGAGTGGCAGCATTTCTGTCAGCAACCACTTTAGAGCGTGGGA 1200  
QY 1201 AAACCTTAGGATGAATATGGGGAATAATGTAAGAGTTCGCGCTTACTGTAGAGGCTATCA 1260  
DB 1201 AAACCTTAGGATGAATATGGGGAATAATGTAAGAGTTCGCGCTTACTGTAGAGGCTATCA 1260  
QY 1261 AAGTAGCGGAAGTGTATGATATAGCAATACACTAAGAAATTAACGTAACCTCTCTCTC 1320  
DB 1261 AAGTAGCGGAAGTGTATGATATAGCAATACACTAAGAAATTAACGTAACCTCTCTCTC 1320  
QY 1321 AACTATTAGTAAATGACGAGACATAAATTTTGGATTAATAAACAATTTAAATCTTCTCT 1380  
DB 1321 AACTATTAGTAAATGACGAGACATAAATTTTGGATTAATAAACAATTTAAATCTTCTCT 1380  
QY 1381 TTCCGTTGAGTTCATTTTCAATTAACCTCCGTTGGATCTTTTCAACGGGAGG 1440  
DB 1381 TTCCGTTGAGTTCATTTTCAATTAACCTCCGTTGGATCTTTTCAACGGGAGG 1440  
QY 1441 TTTTATTGGAAGGTTAAGTATAGTATAGTATCCGATTCAGAGGAATGCTTGAACA 1500  
DB 1441 TTTTATTGGAAGGTTAAGTATAGTATAGTATCCGATTCAGAGGAATGCTTGAACA 1500  
QY 1501 CCTCCGTCACCTAG 1513  
DB 1501 CCTCCGTCACCTAG 1513

RESULT 2  
AAV30255  
ID AAV30255 standard; DNA; 871 BP.  
XX AC AAV30255;  
XX DT 17-OCT-2003 (revised)  
XX DT 27-AUG-2003 (revised)  
XX DT 25-MAR-2003 (revised)  
XX DT 18-AUG-1998 (first entry)  
DE DNA encoding a *Bacillus agaradherens* xylanolytic enzyme.  
XX Xylanolytic enzyme; *Bacillus agaradherens* NCIMB 40482; breakdown;  
KW agricultural waste; alcohol fuel; enzymatic treatment; animal feed;  
KW release; free pentose sugar; dissolving pulp; cellulose; bio-bleaching;  
KW wood pulp; lignocellulosic material; animal feed additive; ss.  
XX *Bacillus agaradherens*.  
XX Key Location/Qualifiers







QY 800 GGTACGTTTCAGTCCCAATGGAACAAATGTTAAACAATATTTTCGTAAGAGTAAAAA 859  
 DB |||||  
 QY 181 GGTACGTTTCAGTCCCAATGGAACAAATGTTAAACAATATTTTCGTAAGAGTAAAAA 240  
 DB |||||  
 QY 860 TTCAATGAAACACAAACACCAACAAAGTTGGTAAACATGTCATAAACTACGAGCCAAAC 919  
 DB |||||  
 QY 241 TTCAATGAAACACAAACACCAACAAAGTTGGTAAACATGTCATAAACTACGAGCCAAAC 300  
 DB |||||  
 QY 920 TTCCAAACCAATGTTGATGCGTATTTATCGGCTATGTTGACGTTGACCTCTTGTGC 979  
 DB |||||  
 QY 301 TTCCAGCAACAGGAATCGTATTTATCGGCTATGTTGACGTTGACCTCTTGTGC 360  
 DB |||||  
 QY 980 GAATATTATTTGTCACAGTTGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGGG 1039  
 DB |||||  
 QY 361 GAATATTATTTGTCATAGTTGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGGG 420  
 DB |||||  
 QY 1040 ACCATCACTGTTGATGCGAGGAACATATGATATCTACGAGCTCTTAGAGTCAATCAACCC 1099  
 DB |||||  
 QY 421 ACCATCACTGTTGATGCGAGGAACATATGATATCTAGAACTCTTAGAGTCAATCAACCC 480  
 DB |||||  
 QY 1100 TCCATTAAGGGGATTCGCACATTTAAACAATTTGGAGTGTTCGAGATCGAAACGCAAG 1159  
 DB |||||  
 QY 481 TCCATTAAGGGGATTCGCACATTTAAACAATTTGGAGTGTTCGAGATCGAAACGCAAG 540  
 DB |||||  
 QY 1160 AGTGGCAGCAATTTCTGTGACGACCACTTTAGAGCGTGGGAACTTAGGATGAATATG 1219  
 DB |||||  
 QY 541 AGTGGCAGCAATTTCTGTGACGACCACTTTAGAGCGTGGGAACTTAGGATGAATATG 600  
 DB |||||  
 QY 1220 GGGAAATGTTGATGCGGCTTACTGTAGAGGCTATCAAAAGTACGCGAAAGTGTAAAT 1279  
 DB |||||  
 QY 601 GGGAAATGTTGATGCGGCTTACTGTAGAGGCTATCAAAAGTACGCGAAAGTGTAAAT 660  
 DB |||||  
 QY 1280 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTAATGACAAG 1339  
 DB |||||  
 QY 661 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTAATGACAAG 720  
 DB |||||  
 QY 1340 AGCATAACTTTGATTAAGAAACAAAT 1363  
 DB |||||  
 QY 721 AGCATAACTTAGTAAGAAACAAAT 744  
 DB |||||

RESULT 5

AAQ92878  
 ID AAQ92878 standard; DNA; 744 BP.  
 XX AC AAQ92878;  
 XX 16-OCT-2003 (revised)  
 DT 12-FEB-1996 (first entry)  
 XX  
 DE Thermostable alkaline endo-1,4-beta-D-xylanase gene.  
 XX  
 KW thermostable alkaline endo-1,4-beta-D-xylanase gene; cloning;  
 KW polymerase chain reaction; Escherichia coli; EC-3.2.1.8; paper; pulp;  
 KW bleaching; ds.  
 XX  
 OS Bacillus sp; 1-43-3 (CBS 672.93).  
 XX  
 XX WO9518219-A1.  
 PN 06-JUL-1995.  
 XX  
 XX 23-DEC-1994; 94WO-EP004312.  
 XX  
 XX 24-DEC-1993; 93EP-00203694.  
 XX  
 XX (KONN ) GIST-BROCADES NV.  
 PA  
 XX Van Solingen P, Williams DP, Iverson S, Farrell RL, Herbes WT;  
 PI Van Der Kleij WA, Herweijer MA, Van Beekhoven RPWC, Quax WJ;  
 PI Goedegebuur F, Jones BE;  
 XX  
 XX WPI; 1995-246385/32.  
 DR

DR P-PSDB; AAR76551.  
 XX Novel xylanase enzyme active at high pH - useful in paper and pulp prodn.  
 PT processes.  
 PS Claim 4; Page 42-43; 54pp; English.  
 XX  
 CC The sequence encodes an alkaline endo-1,4-beta-D-xylanase (G-type) from  
 CC Bacillus sp. 1-43-3. Fragments of the sequence (e.g. the internal  
 CC fragments given in AAQ92878 and AAQ92877) may be amplified by polymerase  
 CC chain reaction, e.g. using primers with sequences AAQ92866, AAQ92867,  
 CC AAQ92868 and AAQ92869. The DNA may be cloned in Escherichia coli using a  
 CC plasmid vector for recombinant xylanase production. The xylanase may be  
 CC used in the paper and pulp industries, where it produces an increase in  
 CC ISO brightness of softwood pulp of at least 1.0 over non-enzymatically  
 CC treated pulp in an ECF pulp bleaching process, at pH 9.0 and 65 deg C.  
 CC The enzyme may be used in production of paper, board and fluff pulp, and  
 CC has low cellulase activity. The increased brightness produced using the  
 CC xylanase allows reduction in the amount of bleaching chemicals used.  
 CC (Updated on 16-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 744 BP; 248 A; 145 C; 159 G; 192 T; 0 U; 0 Other;  
 Query Match 44.5%; Score 673.6; DB 2; Length 744;  
 Best Local Similarity 95.7%; Pred. No. 6.8e-149;  
 Matches 714; Conservative 0; Mismatches 29; Indels 3; Gaps 2;  
 QY 620 ATGACGCAAAAGAAATTCAGGTTGATTTTGTAGCTTTTGTAGTTTGTTCACCTTAACCTTA 679  
 DB 1 ATGACGCAAAAGAAATTCAGGTTGATTTTGTAGCTTTTGTAGTTTGTTCACCTTAACCTTA 57  
 QY 680 CCTGCAGAAATTAATTCAGGCAAAATTCGTACCGCAATTTCCATTTGGCAACCAACGATGCG 739  
 DB 58 CCTGCAAGATTAAGTCAGGCAAAATTCGTACCGCAATTTCCATTTGGCAACCGCGTGGT 117  
 QY 740 TATGATTATGAATTTTGGAAAGATAGCGGTGCTCTGGGCAATGATTTCTCAATCATGCG 799  
 DB 118 TATGATTATGAATTTTGGAAAGATAGCGGTGCTCTGGGCAATGATTTCTCAATCATGCG 177  
 QY 800 GGTACGTTTCAGTCCCAATGGAACAAATGTTAAACAATATTTTCGTAAGAGTAAAAA 859  
 DB 178 GGTACGTTTCAGTCCCAATGGAATTAATGTTAAACAATATTTTCGTAAGAGTAAAAA 237  
 QY 860 TTCAATGAAACACAAACACCAACAAAGTTGGTAAACATGTCATAAACTACGAGCCAAAC 919  
 DB 238 TTCAATGAAACACAAACACCAACAAAGTTGGTAAACATGTCATAAACTACGAGCCAAAC 297  
 QY 920 TTCCAAACCAATGTTGATGCGTATTTATCGGCTATGTTGACCTCTTGTGC 979  
 DB 298 TTCCAGCCAAACGGTAAATCGTATTTATCGGCTATGTTGACCTCTTGTGC 357  
 QY 980 GAATATTATTTGTCACAGTTGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGGG 1039  
 DB 358 GAATATTATTTGTCACAGTTGGGCAACTGGCGTCCACAGGAGCAACGCTTAAGGGG 417  
 QY 1040 ACCATCACTGTTGATGCGAGGAACATATGATATCTACGAGCTCTTAGAGTCAATCAACCC 1099  
 DB 418 ACCATCACTGTTGATGCGAGGAACATATGATATCTAGAACTCTTAGAGTCAATCAACCC 477  
 QY 1100 TCCATTAAGGGGATTCGCACATTTAAACAATTTGGAGTGTTCGAGATCGAAACGCAAG 1159  
 DB 478 TCCATTAAGGGGATTCGCACATTTAAACAATTTGGAGTGTTCGAGATCGAAACGCAAG 537  
 QY 1160 AGTGGCAGCAATTTCTGTGACGACCACTTTAGAGCGTGGGAACTTAGGATGAATATG 1219  
 DB 538 AGTGGCAGCAATTTCTGTGACGACCACTTTAGAGCGTGGGAACTTAGGATGAATATG 597  
 QY 1220 GGGAAATGTTGATGCGGCTTACTGTAGAGGCTATCAAAAGTACGCGAAAGTGTAAAT 1279  
 DB 598 GGGAAATGTTGATGCGGCTTACTGTAGAGGCTATCAAAAGTACGCGAAAGTGTAAAT 657  
 QY 1280 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCAACTATTAGTAATGACAAG 1339  
 DB |||||

658 GTATATAGCAATACACTAAGAATTACGGAAACCCCTCTCTCACTATTAGTAAATACCGAG 717

1340 AGCAATACCTTTGGATATAAAACAAATTA 1365

718 AGCAATACCTTTAGATATAAAACAAATTA 743

# RESULT 6

AAT16101

ID AAT16101 standard; DNA; 663 BP.

XX AAT16101;

DT 16-OCT-2003 (revised)

DT 15-MAY-1996 (first entry)

XX Xylanase gene.

XX Xylanase; thermostable enzyme; paper; pulp; bleaching; feedstuff; baking; ss.

XX Bacillus sp; strain 710/1 (LMG P-14798).

XX Key Location/Qualifiers

FT mat\_peptide 1..663

FT /\*tag= a

FT /EC\_number= "3.2.1.8"

XX AU9525086-A.

XX 08-FEB-1996.

XX 19-JUL-1995; 95AU-00025086.

XX 26-JUL-1994; 94BE-00000706.

XX 17-MAY-1995; 95BE-00000448.

XX (SOLV ) SOLVAY SA.

XX De Buyl E, Lahaye A, Ledoux P, Detroz R;

XX WPI; 1996-117341/13.

XX P-PSDB; AAR92053.

XX Bacillus derived xylanase active over wide pH range - used in treatment of paper pulp, animal feeds and in bakery goods.

XX Claim 30; Page 50-51; 94pp; English.

XX A DNA sequence (AAT16101) coding for a thermostable mature xylanase (AAR92053) was isolated from a gene library of Bacillus sp. 720/1 (LMG P-14798). Sequences were also obtd. (see AAT16102 and AAT16103) for the xylanase precursor and for the complete gene including 5' and 3' untranslated sequences. DNA coding for the mature enzyme may be incorporated into a vector and expressed from either its own promoter or from the Bacillus pumilus PRL B12 promoter (AAQ73996), and used for prodn. of recombinant xylanase in transformed hosts, pref. Bacillus licheniformis or B. pumilus. The enzyme is useful in the paper-pulp, animal feed and baking industries. (Updated on 16-OCT-2003 to standardise OS field)

XX Sequence 663 BP; 220 A; 131 C; 146 G; 166 T; 0 U; 0 Other;

XX Query Match 43.8%; Score 663; DB 2; Length 663;

XX Best Local Similarity 100.0%; Pred. No. 2.1e-146; Indels 0; Gaps 0;

XX Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 701 CAAATCGTCACGCAAAITCCATTCGCAACACGATGGCTATGATTGAATTTGGAAA 760

DB 1 CAAATCGTCACGCAAAITCCATTCGCAACACGATGGCTATGATTGAATTTGGAAA 60

QY 761 GATACGGTGGCTCTGGGACATGATTCTCAATCATGGGGTACGTTCAAGTCCCAATGG 820

Db 61 GATACGGTGGCTCTGGGACATGATTCTCAATCATGGGGTACGTTCAAGTCCCAATGG 120

QY 821 AACCAATGTTAAACAATATTATTCCGTAAGAGTAAAAAATTCATGAACAACAACACAC 880

Db 121 AACCAATGTTAAACAATATTATTCCGTAAGAGTAAAAAATTCATGAACAACAACACAC 180

QY 881 CAACAAGTTGGTAACATGTCATATAAATACCGAGCAACTTCCCAACCAAAATGGTAATCGG 940

Db 181 CAACAAGTTGGTAACATGTCATATAAATACCGAGCAACTTCCCAACCAAAATGGTAATCGG 240

QY 941 TATTTATCGTCTATGTTGGACTGTTGACCCCTCTTGTGCAATATTATTGTGACAGT 1000

Db 241 TATTTATCGTCTATGTTGGACTGTTGACCCCTCTTGTGCAATATTATTGTGACAGT 300

QY 1001 TGGGGCAACTGGCTCCACAGGAGCAACGCTTAAGGGGACCATCATCTGTTGATGGAGGA 1060

Db 301 TGGGGCAACTGGCTCCACAGGAGCAACGCTTAAGGGGACCATCATCTGTTGATGGAGGA 360

QY 1061 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAGGGGATTCGCCACA 1120

Db 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAGGGGATTCGCCACA 420

QY 1121 TTTAAACAATATTGGAGTGTTCGAAAGATCGAAACGACGAGTGGCAGCAATTTCTGTGACG 1180

Db 421 TTTAAACAATATTGGAGTGTTCGAAAGATCGAAACGACGAGTGGCAGCAATTTCTGTGACG 480

QY 1181 AACCACTTTAGAGCGTGGGAAACCTTAGGGGATGAATATGGGAAATCTATGAAGTCCGC 1240

Db 481 AACCACTTTAGAGCGTGGGAAACCTTAGGGGATGAATATGGGAAATCTATGAAGTCCGC 540

QY 1241 CTTACTGTAGAAGGCTATCAAAGTAGCGGAAGTCTTAATGTATATAGCAATACACTAAGA 1300

Db 541 CTTACTGTAGAAGGCTATCAAAGTAGCGGAAGTCTTAATGTATATAGCAATACACTAAGA 600

QY 1301 ATTAACGGTAACCCCTCTCTCACTATTAGTAATGACGAGACATACTTTGGATAAAAAC 1360

Db 601 ATTAACGGTAACCCCTCTCTCACTATTAGTAATGACGAGACATACTTTGGATAAAAAC 660

QY 1361 AAT 1363

Db 661 AAT 663

## RESULT 7

AAV13074

ID AAV13074 standard; DNA; 747 BP.

XX AAV13074;

DT 19-MAY-1998 (first entry)

DE Xylanase activity positive clone DNA SEQ ID NO:11.

KW Bacillus sp; xylanase; glycosyl hydrolase family 11; isolation; microorganism; identification; hybrid DNA; ss.

OS Synthetic.

OS Bacillus sp.

XX Key Location/Qualifiers

FT CDS 1..747

FT /\*tag= a

FT /product= "positive clone for xylanase activity"

XX W09743409-A2.

XX 20-NOV-1997.

XX 12-MAY-1997; 97WO-DK000216.

XX 10-MAY-1996; 96DK-00000562.

XX (NOVO ) NOVO-NORDISK AS.

XX Dalboe H, Diderichsen B, Sandal T, Kauppinen S;  
 XX WPI; 1998-008878/01.  
 XX Isolating novel DNA sequences from microorganisms - without the need for  
 XX culturing the microorganism.  
 XX Example 1; Page 35; 72pp; English.  
 XX The present sequence represents a positive clone for xylanase activity  
 XX from an example of the present invention. The present invention describes  
 XX a novel method for providing a novel DNA sequence encoding a polypeptide  
 XX from a microorganism with an activity of interest. The method comprises:  
 XX (i) PCR amplification of the DNA with PCR primers with homology to (a)  
 XX known gene(s) encoding a polypeptide with an activity of interest; (ii)  
 XX linking the obtained PCR product of a 5' structural gene sequence and a  
 XX 3' structural gene sequence; (iii) expressing the resulting hybrid DNA  
 XX sequence; (iv) screening for hybrid DNA sequences encoding a polypeptide  
 XX with the activity of interest or a related activity; and (v) isolating  
 XX the hybrid DNA sequence identified in step (iv). This method provides for  
 XX identification and isolation of sequences from microorganisms without  
 XX having to cultivate and isolate the microorganism  
 XX  
 XX Sequence 747 BP; 236 A; 168 C; 154 G; 189 T; 0 U; 0 Other;  
 SQ  
 Query Match 39.0%; Score 590.2; DB 2; Length 747;  
 Best Local Similarity 86.9%; Pred. No. 3.2e-129;  
 Matches 649; Conservative 0; Mismatches 98; Indels 0; Gaps 0;  
 QY 620 ATGAGCAAAAGAAATGACGTTGATTTAGCTTTTGTAGTTTGTGCTACCTTAACCTTA 679  
 DB 1 ATGAGCAAAAGAAATGACGTTGATTTAGCTTTTGTAGTTTGTGCTACCTTAACCTTA 60  
 QY 680 CTGCGAAGAAATTAATTCAGGCAAAATCGTCACGCAATTCATTCGCAACCAACGATGCG 739  
 DB 61 CTGCGAAGAAATTAATTCAGGCAAAATCGTCACGCAATTCATTCGCAACCAACGATGCG 120  
 QY 740 TATGATTATGAATTTGGAAAGATAGCGGTGCTCTGGGACAAATGATTCATCATGCG 799  
 DB 121 TATGATTATGAATTTGGAAAGATAGCGGTGCTCTGGGACAAATGATTCATCATGCG 180  
 QY 800 GGTACGTTGAGTCCCAATGGAACATGTTAAACAATATTTCCGTAAGGTAAAAA 859  
 DB 181 GGTACGTTGAGTCCCAATGGAACATGTTAAACAATATTTCCGTAAGGTAAAAA 240  
 QY 860 TTCATGAAACAAACACACCAACCAAGTTGGTAAATGTCATGTCATGTCATGTCATGTC 919  
 DB 241 TTCATGAAACAAACACACCAACCAAGTTGGTAAATGTCATGTCATGTCATGTCATGTC 300  
 QY 920 TTCACCAACCAATGGAATGATTTATGCGTCTATGTTGGTGGTGGTGGTGGTGGTGGTGG 979  
 DB 301 TTCACCAACCAATGGAATGATTTATGCGTCTATGTTGGTGGTGGTGGTGGTGGTGGTGG 360  
 QY 980 GAATATTTATTCGACAGTTGGGCACTGGGCACTGGGCACTGGGCACTGGGCACTGGGCACT 1039  
 DB 361 GAATATTTATTCGACAGTTGGGCACTGGGCACTGGGCACTGGGCACTGGGCACTGGGCACT 420  
 QY 1040 ACCATCACTGTCATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 1099  
 DB 421 ACCATCACTGTCATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480  
 QY 1100 TCCATTAAGGGGATGTCACATTTAAACAATATGAGAGTGTTCGAAGATCGAAACGACG 1159  
 DB 481 TCTATTTCAGGGCACCGGCACCTTCAATCAGTACTGTTGATTCGACAGCAAGCGGAC 540  
 QY 1160 AGTGGACAGATTTCTGTCACACCACTTTAGAGCGTGGGCACTTAGGAGTGAATATG 1219  
 DB 541 AGCGGCATGTCATACGCAAAACCACTTTAATGCTCGGGCTGCTCTTGGCATGAATG 600  
 QY 1220 GGGAAATGATGAGTGGCGCTTACTGTAGAGGCTTATCAAAAGTACGCGAAGTCTAAT 1279  
 DB 601 GGTGCATTCATTTACAGATCCTCGTTACTGAGGGCTACCAATCTACCGAAGTCTAAT 660

QY 1280 GTATATAGCAATACACTAAGAAATTAACGGTAACCTCTCTCACTATTAGTAATGACGAG 1339  
 DB 661 GTATATAGCAATACACTAAGAAATTAACGGTAACCTCTCTCACTATTAGTAATGACGAG 720  
 QY 1340 AGCATAACTTTGGATATAAAACAATTA 1366  
 DB 721 AGCATAACTCTAGATAAAACAATTA 747  
 RESULT 8  
 ADJ34947  
 ID ADJ34947 standard; DNA; 1068 BP.  
 AC ADJ34947;  
 XX  
 DT 22-APR-2004 (first entry)  
 XX  
 DE DNA encoding xylanase from an environmental sample seq id 163.  
 XX  
 KW antibacterial; fungicide; thermostable xylanase activity;  
 KW dough conditioning; beverage production; nutritional supplement;  
 KW animal feed; lignin reduction; wood product; xylan; bacterial infection;  
 KW fungal infection; coccidiosis; gene; ds.  
 XX  
 OS Unidentified.  
 XX  
 XX WO2003106654-A2.  
 PN  
 XX 24-DEC-2003.  
 PD  
 XX 16-JUN-2003; 2003WO-US019153.  
 PF  
 XX 14-JUN-2002; 2002US-0389299P.  
 PR  
 XX (DIVE-) DIVERSA CORP.  
 PA  
 XX Steer B, Callen W, Healey S, Hazlewood G, Wu D, Blum D;  
 PI Esteghlalian A;  
 XX  
 XX WPI; 2004-099016/10.  
 DR P-PSDB; ADJ34948.  
 XX  
 XX Novel xylanase recombinant polypeptide useful for improving textile  
 XX texture, treating paper, eliminating microorganisms.  
 PT  
 PT  
 XX Claim 1; SEQ ID NO 163; 570pp; English.  
 PS  
 CC The invention describes an isolated or recombinant polypeptide (I),  
 CC having 50% or more identity to 190 300-1200 residue amino acid sequences  
 CC (SI), given in the specification, over a region of 100 or more residues  
 CC and the polypeptide as thermostable xylanase activity. (I) is useful for:  
 CC dough conditioning; beverage production; as a nutritional supplement in  
 CC animal feed; reducing lignin in a wood or a wood product; and for  
 CC eliminating and protecting animals from a microorganism comprising xylan.  
 CC The polynucleotide (II) encoding (I) is useful for amplifying nucleic  
 CC acid encoding a polypeptide having a xylanase activity which involves  
 CC amplification of a template nucleic acid with a primer pair capable of  
 CC amplifying (II) or its subsequence. (I) is useful for treating and  
 CC preventing bacterial infection and fungal infection e.g. coccidiosis.  
 CC This sequence encodes xylanase protein isolated from an environmental  
 CC sample.  
 XX  
 SQ Sequence 1068 BP; 303 A; 249 C; 279 G; 237 T; 0 U; 0 Other;  
 Query Match 26.0%; Score 393; DB 12; Length 1068;  
 Best Local Similarity 75.2%; Pred. No. 1.1e-82;  
 Matches 518; Conservative 0; Mismatches 165; Indels 6; Gaps 2;  
 QY 620 ATGAGCAAAAGAAATGACGTTGATTTAGCTTTTGTAGTTTGTGCTACCTTAACCTTA 679  
 DB 1 ATGAGCAAAAGAAATGACGTTGATTTAGCTTTTGTAGTTTGTGCTACCTTAACCTTA 60

QY 680 CCTGCAGAAATATTCAGGCAAAATCGTCACGCAAAATTCATTGGCAACACCATGATGC 739  
Db |||||  
QY 61 CCTGGG---CAGTGATCGCAGACGATCACGCAATTCGGTCGTACGATACGGT 117  
Db |||||  
QY 740 TATGATATGAATTTTGGAAAGATAGCGGTGGCTCTGGGACAATGATTCTCAATCATGCG 799  
Db |||||  
QY 118 TATGACTATGAATCTTGAAGGACAGC---GGGAATGGAATCTATGTTCTCGGTAGTGC 174  
QY 800 GGTACGTTTCAGTGCCCAATGGACAAATGTTTACACACATATTTCCGTAAAGTTAAAAA 859  
Db |||||  
QY 175 GGTACGTTTCAGTGCCGAGTGGAGCAATATCAATAATATTTCTGTTCCGTAAGGCAAGAG 234  
QY 860 TTCAATGAACACAAACACACAAAGTGTGTAAATGTCATCAATGTCATCAATGACGCAAC 919  
Db |||||  
QY 235 TTCAATGACGCGACGACCCATCAGCAATTTGGAACATTTCCATTAACCTATGTCGCCACC 294  
QY 920 TTCCAAACCAATGGTAATCGGTATTTATGCGTCTATGTTGACCTTGTGACCTCTTGTGC 979  
Db |||||  
QY 295 TACCAACCAATGGCAATTCGTATTTAAACGGTCTATGCTGACGGTTGACCCCTCGTC 354  
QY 980 GAATATTATTTGTCGACAGTTGGGCACTGGCGTCCACAGGACCAACGCTTAAGGG 1039  
Db |||||  
QY 355 GAATATTATGTCGATAGCTGGGCGACGCTGGCGTCCGCTGGAGCATCGCCAAAGGG 414  
QY 1040 ACCATCACTGTTGATGGAGCAATATGATATCTACGACCTTTAGAGTCAATCAACCC 1099  
Db |||||  
QY 415 ACTGTTAAGTTGACGGAGNACGTATGACATTTATGACAACTCGTGTCAACAGCT 474  
QY 1100 TCCATTAGGGGATTCGCACATTTAAACAAATTTGAGTGTTCGAAGATTCGAAGCGACG 1159  
Db |||||  
QY 475 TCCATTAAAGGACGCGCAACCTTCAAGCAGTATTTGAGTGTTCGGACGCTCAAAACGGACG 534  
QY 1160 AGTGGCAGATTTCTGTACGCAACCACTTTAGAGCGTGGGAACTTAGGGATGATATG 1219  
Db |||||  
QY 535 ACAGGAAACCATTTCTGTAAGCGAGCACCTTTAAGGCTGGGAAATTTGGGATGACCATG 594  
QY 1220 GGGAAATGTTATGAAGTCGCGCTTACTGTAGAAGCTATCAAGTAGCGGAGTGTCTAAT 1279  
Db |||||  
QY 595 GCGAAGATGATGAAGTCGCGCTTACGGTTGAGGCTTCAAGCAGTGGGAGCGCTAAT 654  
QY 1280 GTATATGCAATACATPAGAAATTAACGG 1308  
Db |||||  
QY 655 GTGTATAGCCATACATGACGATCGCGG 683

## RESULT 9

AAQ80923

ID AAQ80923 standard; DNA; 1022 BP.

AC AAQ80923;

DT 02-AUG-1995 (first entry)

DE B. pumilus xylanase gene.

KW Xylanase; endo-1,4-beta-D-xylanase; wood pulp; pulping; biobleaching;  
KW bleaching; Bacillus licheniformis; ss.  
OS Bacillus pumilus.

Key Location/Qualifiers  
FH 5'UTR 1..185  
FT /\*tag= a  
FT /note= "claim 12"  
FT misc\_difference 107  
FT /\*tag= b  
FT /note= "base n at position 107 is not identified in the  
FT specification"  
FT CDS 186..869  
FT /\*tag= c  
FT sig\_peptide 186..266  
FT /\*tag= d  
FT /note= "claim 11"

mat\_peptide 267..866  
/\*tag= e  
/EC\_number= "3.2.1.8"  
/note= "claim 10"

GB2279955-A.  
18-JAN-1995.  
15-JUL-1993; 93GB-00014780.  
15-JUL-1993; 93GB-00014780.  
(SOLV ) SOLVAY & CIE.  
Amory A, Andre C, De Buyl E, Detroz R, Lahaye A, Ledoux P;  
WPI; 1995-039214/06.  
P-PSDB; AAR68849.

Purified xylanase from Bacillus pumilus PRL B12 - esp. produced in  
transformed Bacillus licheniformis, and related DNA, vectors, etc., used  
for pre-treatment of wood pulp to reduce chlorine or ozone consumption in  
subsequent bleaching.

Claim 13; Fig 1a-1b; 97pp; English.

A Bacillus pumilus PTL B12 (ATCC 55443) gene library was screened for  
recombinant plasmids carrying the xylanase gene. A chromosomal fragment  
obtained from isolate pBPX1 was subcloned and expressed in Escherichia  
coli JM109. The sequence of the xylanase gene (as a 1022 bp Sau3AI  
fragment) carried by a selected transformant is given in AAQ80923

Sequence 1022 BP; 329 A; 197 C; 229 G; 266 T; 0 U; 1 Other;

Query Match 18.7%; Score 282.8; DB 2; Length 1022;  
Best Local Similarity 62.1%; Pred. No. 1.1e-56;  
Matches 481; Conservative 0; Mismatches 287; Indels 6; Gaps 2;

QY 599 AATTTAAAGGAGGATGCTTAATGAGACAAAGAAATGAGCTGATGATTTAGCTTTTGA 658  
Db |||||  
QY 165 AAAAAGGAGGAGGATGGAATTTGAAAAGATTTAGGCTGTTGTTGTGATGTT 224  
QY 659 GTTTGTTTTCACCTAACCTTACCTGACAGAAATTAATTCAGGCACAAATCGTCACGACAAT 718  
Db |||||  
QY 225 ATTGATTTGCTGACACTGACGCTGTGCGGCTCATGCGGAACGATTTATGATAT 284  
QY 719 TCCATTGCGCAACACGATGGCTATGATTATGAAATTTTGGAAAGATAGCGGTCTGGG 778  
Db |||||  
QY 285 AGGATAGGGACACACAGCGGATACGATTTTGAATTTATGGAAGGATTAC---GGAATACC 341  
QY 779 ACATGATTTCTATCATCATGCGGTACGTTGAGTCCCAATGGAACATGTTAACACATA 838  
Db |||||  
QY 342 TCGATGACACTCAATAACGGCGGGCATTTAGTCAAGCTGGAACATATTTGGAATGCC 401  
QY 839 TTATTCGTAAGGTAAAAAATTCATGAAACACAAACACACCAAGTTGGTAAACATG 898  
Db |||||  
QY 402 TTATTTGAAAGGAAGAAGTTGATTCCTAAACTCATCATCAACTTTGGCAATC 461  
QY 899 TCCATAAAGTACGAGGCAACCTTCCAAACCAATTTGTAATCGGTATTTATCGTCTATGTT 958  
Db |||||  
QY 462 TCCATCAACTACACAGCGAGCTTTAACCCGGGGGAAATTCCTATTATTTATGTCATGCG 521  
QY 959 TGGACTGTTGACCTCTTGTGCAATATATATTTGTGACAGTTGGGCAACTGGCGTCCA 1018  
Db |||||  
QY 522 TGGACAAATCTCATTAGCTGAATACTACATTTGTTGAGTCATGCGGCAATATCGTCCA 581  
QY 1019 CCAGGAGCAACGCTAAGGGGACCATCACCTGTTGATGGAGGAAACATATGATATACGAG 1078  
Db |||||  
QY 582 ACAGG---ACGTATAAGGATCATTTTATGCCGATGGAGGCACATATGACATATGAA 638  
QY 1079 ACTCTTAGAGTCAATCAACCCCTCAATTAAGGGATTTGCCACATTTAAACAAATTTGAGT 1138  
Db |||||

Db 639 AGCTCCGTCATCAGCCTTCATCATTTGGAGCGCTACCTTCAACAATATTGGAGT 698  
 QY 1139 GTTGAAGATCGAAGCGACGAGTGGCAGGATTTCTGTCAGCAACACTTTAGACGGTGG 1198  
 Db 699 GTACGTCACAAACAAAGCGACAGCGGACCGTCTCGCTCAGTGAGCATTTTAAAAAATGG 758  
 QY 1199 GAAACTTAGGATGAATATGCGGAAATGTATGAGTGGCGCTTACTGTAGAAGGCTAT 1258  
 Db 759 GAAAGCTTAGGATGCCATTTGGGAAAAATGTATGAAACAGCATTAACGTAGAGGCTAC 818  
 QY 1259 CAAAGTAGCGGAGTGTCTAATGTATATAGCAATACACTAAGAAATTAACGGTAACCTCTC 1318  
 Db 819 CGAAGCAACGAAAGTGGGAATGTATGACGAATCAGCTGATGATTCGATANAAGCATATG 878  
 QY 1319 TCAACTTATTAGTAATGACGAGACATACTTTGGATAAAAACAATTAATAATCC 1372  
 Db 879 AAAAAGCCAGCAAAAAATGGCTGGCTTTTCTATGATAAATTTTCAACTTCC 932

## RESULT 10

ADJ35051

ID ADJ35051 standard; DNA; 1956 BP.

XX

AC ADJ35051;

XX

22-APR-2004 (first entry)

XX

DNA encoding xylanase from an environmental sample seq id 267.

XX

antibacterial; fungicide; thermostable xylanase activity;

KW

dough conditioning; beverage production; nutritional supplement;

KW

animal feed; lignin reduction; wood product; xylan; bacterial infection;

KW

fungal infection; coccidiosis; gene; ds.

XX

Unidentified.

OS

XX

WO2003106654-A2.

XX

24-DEC-2003.

XX

16-JUN-2003; 2003WO-US019153.

XX

14-JUN-2002; 2002US-0389299P.

XX

(DIVE-) DIVERSA CORP.

XX

Steer B, Callen W, Healey S, Hazlewood G, Wu D, Blum D;

XX

Esteghlalian A;

PI

WPI; 2004-099016/10.

XX

P-PSDB; ADJ35052.

XX

Novel xylanase recombinant polypeptide useful for improving textile

PT

texture, treating paper, eliminating microorganisms.

PT

Claim 1; SEQ ID NO 267; 570pp; English.

XX

The invention describes an isolated or recombinant polypeptide (I),

XX

having 50% or more identity to 190 300-1200 residue amino acid sequences

CC

(S1), given in the specification, over a region of 100 or more residues

CC

and the polypeptide as thermostable xylanase activity. (I) is useful for:

CC

dough conditioning; beverage production; as a nutritional supplement in

CC

animal feed; reducing lignin in a wood or a wood product; and for

CC

eliminating and protecting animals from a microorganism comprising xylan.

CC

The polynucleotide (II) encoding (I) is useful for amplifying nucleic

CC

acid encoding a polypeptide having a xylanase activity which involves

CC

amplification of a template nucleic acid with a primer pair capable of

CC

amplifying (II) or its subsequence. (I) is useful for treating and

CC

preventing bacterial infection and fungal infection e.g. coccidiosis.

CC

This sequence encodes xylanase protein isolated from an environmental

CC

sample.

CC

Sequence 1956 BP; 614 A; 384 C; 423 G; 535 T; 0 U; 0 Other;

XX

SQ

Query Match 18.3%; Score 276.6; DB 12; Length 1956;  
 Best Local Similarity 66.2%; Pred. No. 3.8e-55;  
 Matches 415; Conservative 0; Mismatches 209; Indels 3; Gaps 1;  
 QY 688 AATAATTCAGGCACAAATCGTCACGCAATTCATTCGCAACACGATGGCTATGATTA 747  
 Db 81 AGTACTCCGCCGGCGAATAATTTACGACAATGAGACAGGCACACATGAGGCTACGACTA 140  
 QY 748 TGAATTTGGAAAGATAGCGGTGCTCTGGGACATGATTCCTCAATCATGGCGGTACGTT 807  
 Db 141 TGAGCTCTCGAAAGA---CTACGGAAATACGATTTATGGAACCTTAACGCGGTGGTACTTT 197  
 QY 808 CAGTCCCAATGGAACAATGTTAAACAATATTTTCGGTAAAGGTAAAAAATTCATGA 867  
 Db 198 TAGTTGTCAATGGAGTAAATATCGGTAATGCACATTTTAGMAAAGGAGAAAAATTAATTC 257  
 QY 868 AACACAAACACACCAACCAAGTTGGTAAATGTCATTAACCTAGGAGGCCAATCTCCAAC 927  
 Db 258 CGACAAAACCTATCAAGAAATTAGGAGATATAGTAGTTGAATATGGCTGTGATTACAATCC 317  
 QY 928 AAATGGTAAATGCTATTTATGCTATGTTGGACTGTTGACCTCTTGTGCAATATTA 987  
 Db 318 AAACCGAAATTCCTATTTGTGTTTACGGTTGGACAAGAAATCCACTGTTGAATATTA 377  
 QY 988 TATTGTGACAGCTTGGGGCAACTGGCGTCCACAGGAGCAACCGCTTAAGGGACCATCAC 1047  
 Db 378 CATTTAGAAAGCTGGGCGAGCTGGGTCACCTGGAGCAACACCCAAAGGAACCATCAC 437  
 QY 1048 TGTGATGGAGGAACATATGATATCTAGAGACTCTTAGAGTCAATCAACCTCCATTA 1107  
 Db 438 AGTGGATGGCGTACTTATGAAATATATGAAACTACCCGGGTAAATCAGCCTTCCATCGA 497  
 QY 1108 GGGGATGGCACATTTAAACAATATTTGGAGTCTTCGACATCGAACGACGAGTGGCAC 1167  
 Db 498 TGGAACTGCGACATTTCCAAACAATATTTGGAGTGTGTGATCAACAGAACAAAGCCG 557  
 QY 1168 GATTCTGTGACCAACCACTTTTAGAGCGTGGGAAAACCTTAGGGATGAATATGGGAAAAT 1227  
 Db 558 AATACTGTCTACTGAAACATTTTAAACAGTGGGAAAGATGGGCATCGAATGGTAAAT 617  
 QY 1228 GTATGAAGTCGCGCTTACTGTAGAAGGCTATCAAGTACGGAAGTCTTAATGTATATAG 1287  
 Db 618 GTATGAAGTTGCTCTTACCGTTGAAGTTATCAGAGCAGTGGGTACGCTAATGTATATA 677  
 QY 1288 CAATACACTAAGAAATTAACGGTAAACCC 1314  
 Db 678 GAATGAATCAGAATAGGTGCAAAATCC 704

## RESULT 11

ADJ34969

ID ADJ34969 standard; DNA; 684 BP.

XX

AC ADJ34969;

XX

22-APR-2004 (first entry)

XX

DNA encoding xylanase from an environmental sample seq id 185.

XX

antibacterial; fungicide; thermostable xylanase activity;

KW

dough conditioning; beverage production; nutritional supplement;

KW

animal feed; lignin reduction; wood product; xylan; bacterial infection;

KW

fungal infection; coccidiosis; gene; ds.

XX

Unidentified.

OS

XX

WO2003106654-A2.

XX

24-DEC-2003.

XX

16-JUN-2003; 2003WO-US019153.

XX



PR 14-JUN-2002; 2002US-0389299P.  
XX (DIVE-) DIVERSA CORP.  
XX Steer B, Callen W, Healey S, Hazlewood G, Wu D, Blum D;  
PI Esteghlalian A;  
XX WPI; 2004-099016/10.  
DR P-PSDB; ADJ34970.  
XX Novel xylanase recombinant polypeptide useful for improving textile  
PT texture, treating paper, eliminating microorganisms.  
XX Claim 1; SEQ ID NO 185; 570pp; English.  
XX The invention describes an isolated or recombinant polypeptide (I),  
CC having 50% or more identity to 190 300-1200 residue amino acid sequences  
CC (SI), given in the specification, over a region of 100 or more residues  
CC and the polypeptide as thermostable xylanase activity. (I) is useful for:  
CC dough conditioning; beverage production; as a nutritional supplement in  
CC animal feed; reducing lignin in a wood or a wood product; and for  
CC eliminating and protecting animals from a microorganism comprising xylan.  
CC The polynucleotide (II) encoding (I) is useful for amplifying nucleic  
CC acid encoding a polypeptide having a xylanase activity which involves  
CC amplification of a template nucleic acid with a primer pair capable of  
CC amplifying (II) or its subsequence. (I) is useful for treating and  
CC preventing bacterial infection and fungal infection e.g. coccidiosis.  
CC This sequence encodes xylanase protein isolated from an environmental  
CC sample.  
XX  
XX  
SQ Sequence 684 BP; 221 A; 129 C; 158 G; 176 T; 0 U; 0 Other;  
Query Match 18.1%; Score 274.4; DB 12; Length 684;  
Best Local Similarity 64.6%; Pred. No. 9.5e-55;  
Matches 442; Conservative 0; Mismatches 236; Indels 6; Gaps 2;  
QY 620' ATGACACAAAAGAAATTGACGTTGATTTTATGACCTTTTATGTTTGTTCACCTAACCTTA 679  
DB 1 ATGAATTTGAAAAGATTGAGCGCTGTTTGTGATGTGATTGGATTGTGCTGACACTG 60  
QY 680 CTTGAGAAATTAATTCAGGACCAATTCGTACCGCAATTCATTTGGCAACACGATGGC 739  
DB 61. ACGGCTGTGCCAGCTCATGCGGAAACGATTTATGATAATAGGATAGGACACACAGCGGA 120  
QY 740 TATGATTTATGAATTTTGGAAAGATAGCGTGCTCTGGGACATATGATTCCTCAATCATGCG 799  
DB 121 TACGATTTTGAATTAAGAGATTACGGAATACCTCG---ATGACACTCAATACCGC 177  
QY 800 GGTACGTTTCAGTCCCAATGGAACAAATGTTTAAACAATATTTATTCGTAAGGTAAAAA 859  
DB 178 GGGGCATTTAGTGCAAGCTGGAACAATATTGGAAATGCTTATTTTCGAAAAGGAAGAAG 237  
QY 860 TTCAATGAAACACAAACACCAACAGATTTGGTAAACATGTCATAAATCAGGAGCCAAC 919  
DB 238 TTGATTTCCACTAAAACATCATCACTTGGCAACATCTCATCACTCAACACGACCC 297  
QY 920 TTCCAAACCAATGTAATGCGTATTTATGCGTCTATGTTGACGTTGTGACCTCTTCTC 979  
DB 298 TTTACCCGGGGGAATTCCTATTTATGTTCTAAGCTGAGACAAATCTCCATTAGCT 357  
QY 980 GAATATTATTTGTCACAGTTTGGGGCAACTGGCGTCCACGAGGACCAACGCTTAAGGGG 1039  
DB 358 GAATACTACATTTGTTGAGTCATGCGGACATATCTGTCACACAGG---AACGTATAAAGGA 414  
QY 1040 ACCATCATGTTGATGGAGGAACATATGATATCTAGAGACTCTTAGAGTCAATCAACCC 1099  
DB 415 TCATTTTATGCGATGGAGGACATATGACATATATGAACGCTCCGTGTCAATCAGCCT 474  
QY 1100 TCCATTATAGGGGATTCACACATTTAAACAATTTGAGTGTTCGAAGATCGAAACGACG 1159  
DB 475 TCTATCATTTGAGAGCGCTACCTTCAACANATTTGAGTGTACGTCAACAAACGACACA 534  
QY 1160 AGTGGCACGATTTCTGTGACGAAACCACTTTAGAGCGTGGGAAACTTAGGGATGAATATG 1219

DB 535 AGCGAACTGTTCCGTCAGTGAGCATTTTAAAAATGGGAAAGCTTAGGCATGCCAATG 594  
QY 1220 GGGAAATGTATGAAGTCGCGCTTACTGTAGAGGCTATCAAGTAGCGGAGTGCTAAT 1279  
DB 595 GGAATAATGTATGAACAGCATTTAACTGTAGAGGCTACCGAAGCAACGGAAGTGCGAAT 654  
QY 1280 GTATATACCAATACACTAAGAATT 1303  
DB 655 GTCATGCGAATCAGCTGATGATT 678  
RESULT 12  
AAZ51821  
ID AAZ51821 standard; DNA; 2364 BP.  
XX  
AC AAZ51821;  
XX  
DT 04-JUL-2000 (first entry)  
XX Clostridium stercorarium xylanase A DNA.  
XX Xylanase A; XynA; cellulosome; feruloyl esterase; phenolic acid esterase;  
KW thermostable; ferulic acid; wheat bran; agricultural byproduct; treat;  
KW grass; paper and pulp industry; feed processing; food additive;  
KW plant cell wall material; degradation; ds.  
XX Clostridium stercorarium.  
OS  
XX  
FH Key Location/Qualifiers  
CDS 440..1978  
FT /\*tag= e  
FT /product= "Xylanase A"  
XX WO200014243-A1.  
XX 16-MAR-2000. 99WO-US020304.  
XX 03-SEP-1999; 99WO-US020304.  
XX 04-SEP-1998; 98US-0099136P.  
XX (UYGE-) UNIV GEORGIA RES FOUND INC.  
XX Blum DL, Kataeva I, Li X, Ljungdahl LG;  
XX WPI; 2000-256991/22.  
DR P-PSDB; AAY70523.  
XX New recombinant DNA molecule comprising a sequence encoding feruloyl  
PT esterase protein, useful for treating grasses and other plant materials  
PT used in pulp and paper industries, feed processing and food additives.  
XX  
XX Disclosure; Page 97-99; 105pp; English.  
CC The present sequence is a DNA encoding xylanase A (XynA) from Clostridium  
CC stercorarium. XynA contains family VI cellulose binding domain (CBD)  
CC which is homologous to the CBD of xylanase Z (XynZ) of Clostridium  
CC thermocellum. XynZ is an enzymatic component of C. thermocellum  
CC cellulosome and has a multi-domain structure which includes a dockerin  
CC domain, a catalytic xylanase domain, a family VI cellulose binding domain  
CC and a domain of unknown function. The unknown domain in the N-terminal  
CC region of XynZ has been found to contain feruloyl (phenolic acid)  
CC esterase which is involved in the degradation of plant cell wall  
CC material. The novel feruloyl esterase is thermostable, easy to purify,  
CC has high temperature optima and stable over a wide pH range. The enzyme  
CC is used for producing ferulic acid from wheat bran or agricultural  
CC byproducts, treating grasses or other plant materials used in the pulp  
CC and paper industries, in feed processing and as a food additive  
XX  
SQ Sequence 2364 BP; 791 A; 428 C; 458 G; 687 T; 0 U; 0 Other;  
Query Match 16.9%; Score 255.6; DB 3; Length 2364;



WO9736995-A2.  
 09-OCT-1997.  
 27-MAR-1997; 97WO-NZ000042.  
 29-MAR-1996; 96NZ-0028296.  
 (PACI-) PACIFIC ENZYMES LTD.  
 Bergquist PL, Gibbs MD, Morris D;  
 WPI; 1997-503090/46.  
 P-PSDB; AAW30267.  
 Dictyoglomus thermophilum xylanase gene - used for producing thermally  
 stable enzymes for the bleaching of cellulase products, especially paper  
 pulp.  
 Disclosure; Fig 3; 34pp; English.  
 The present sequence represents a RT45B.1 xynB gene and encodes a xylanase  
 enzyme. The present sequence is derived from a gene contained within  
 Dictyoglomus thermophilum strain RT45B.1. The xylanase is contained  
 within the family of enzymes known as G-Xylanases, and has beta -1,4-  
 xylanase activity. The enzyme has high activity and high thermal  
 stability with optimum activity at 85 degrees Celcius and pH 6.5. The  
 xylanase enzyme is used for the biological bleaching of cellulose  
 products, especially paper pulp. Use of the enzyme ensures that waste  
 streams from the biological bleaching will include less toxic material  
 Sequence 1190 BP; 386 A; 197 C; 277 G; 330 T; 0 U; 0 Other;

RESULT 13	
AAT90972	
ID	AAT90972 standard; DNA; 1190 BP.
XX	
AC	AAT90972;
XX	
DT	22-MAY-1998 (first entry)
XX	
DE	Nucleotide sequence encoding an enzyme with xylanase activity.
XX	
KW	xynB gene; xylanase; enzyme; Dictyoglomus thermophilum strain Rt46B.1;
KW	G-xylanase; beta -1,4-xylanase activity; activity; thermal stability;
KW	biological bleaching; cellulose product; paper pulp; ss.
XX	
OS	Dictyoglomus thermophilum.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	67..1149
FT	/tag= a
FT	/product= "xylanase"
FT	67..138
FT	/tag= b
FT	139..1146
FT	mat_peptide
FT	/tag= c
XX	

  

QY	774	CTGGGCAANTATTCTCAATCATGCGGTAGCTTCAGTGCACAATGGAAACAATGTTAACA	833
DG			
DG	218	ATAACACCATGACTGTATACACACAAGGAAGGTTTAGCTGTAGTGGAGCAATATAACA	277
QY	834	ACATATTATTCGTTAAAGGTAAAAAATTCAATGAAAACACAAACACACCAACAAGTTGGTA	893
DG			
DG	278	ATGCATTATTGACACAGGTAGAAGTACAA-----CCAAAACCTGGCAGTCATTAGGCA	331
QY	894	ACATGTCCATAAACTACGGAGCCAACCTTCCAACCAAATGGTAATGCGTATTATGCGTCT	953
DG			
DG	332	CTATTAGNATCACCTACTCAGCCACATATAATCTTAATGGTAACTCCTACTTATGTATCT	391
QY	954	ATGTTTGGACTGTTTGACCCCTCTTGTGCAATATTAATTGTCGACAGTTTGGGGCACTGCC	1013
DG			
DG	392	ATGTTTGGTCTACTAATCCTTTAGTAGAGTTTTACATTGTAGAAAAGTTGGGGTAATTGGC	451
QY	1014	GTCCACGAGGACCAAGCCCTTAAGGGGACCATCACCTGTTGTAGGAGGAACATATGATATCT	1073
DG			
DG	452	GTCCACGAGTGCAACCTCTCTTTGGACAGGTACTATATGACGGTGGTACTATGACATTT	511
QY	1074	ACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAAGGGATTTGCCACATTTAAACAATATT	1133
DG			
DG	512	ACAGAACTACCCGTGTAAATCAGCCATCTATTGTGCGGTACAGCTACTTTTGTATCAATATT	571



XX WPI: 2004-099016/10.  
DR P-PSDB; ADJ35012.  
XX  
PT Novel xylanase recombinant polypeptide useful for improving textile  
PT texture, treating paper, eliminating microorganisms.  
XX  
XX  
PS Claim 1: SEQ ID NO 227; 570pp; English.  
XX  
CC The invention describes an isolated or recombinant polypeptide (I),  
CC having 50% or more identity to 190 300-1200 residue amino acid sequences  
CC (S1), given in the specification, over a region of 100 or more residues  
CC and the polypeptide as thermostable xylanase activity. (I) is useful for:  
CC dough conditioning; beverage production; as a nutritional supplement in  
CC animal feed; reducing lignin in a wood or a wood product; and for  
CC eliminating and protecting animals from a microorganism comprising xylan.  
CC The polynucleotide (II) encoding (I) is useful for amplifying nucleic  
CC acid encoding a polypeptide having a xylanase activity which involves  
CC amplification of a template nucleic acid with a primer pair capable of  
CC amplifying (II) or its subsequence. (I) is useful for treating and  
CC preventing bacterial infection and fungal infection e.g. coccidiosis.  
CC This sequence encodes xylanase protein isolated from an environmental  
CC sample.  
XX  
SQ Sequence 747 BP; 174 A; 222 C; 244 G; 107 T; 0 U; 0 Other;  
Query Match 11.8%; Score 177.8; DB 12; Length 747;  
Best Local Similarity 66.1%; Pred. No. 5.9e-32;  
Matches 257; Conservative 0; Mismatches 132; Indels 0; Gaps 0;  
QY 893 AACATGTCATTAACACGAGCGCAACTTCCACCAAAATGGTAATGCGTATTATGCGTC 952  
DB 340 AATCAGACGGTGACCTACAGGCAAACTACCGCCGACCGCAATTCATCCTGTGCGTA 399  
QY 953 TATGGTTGGACTGTTGACCCCTCTTGTGCAATATTATATTGTCGACAGTTGGGGCAACTGG 1012  
DB 400 TATGGATGCGCGCAAAACCCCTCGTGAATACTACATCGTCGATAGTGGGGCAGCTGG 459  
QY 1013 CGTCCACGAGGACGCGCTAAGGGGACCATCCTGTTGATGGAGGACATATGATATC 1072  
DB 460 CGCCCGCGGGGGAACGTCATGGGACCGTCACCGGACGGCGGCACCTACGACATC 519  
QY 1073 TACGAGACTCTTAGAGTCAATCAACCTCCATTAAAGGGATTGCCACATTTAAACAATAT 1132  
DB 520 TACCGCACCCAGCGCGTCAACAGCCTTCCATCGAAGGCACCAAGACCTTCTATCAATAC 579  
QY 1133 TGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGCATTTCTGTGAGCAACCACTTTAGA 1192  
DB 580 TGGAGCGTTTCGCACTCAGAAGCGCACGAGCGGAACGATCACGGTTGCGGCTCCTTCGAC 639  
QY 1193 GCSTGGGAACCTTAGGGATGAATNTGGGAAAAATGTATGAAGTCGGCTTACTGTAGAA 1252  
DB 640 GCCTGGGCGGAGGAGGATGAACATGGGGAGTCTGTACGAGGTGTCGATGACCGCGAG 699  
QY 1253 GGCTATCAAGTAGCGGAAGTGCTAATGT 1281  
DB 700 GGCTATCAAGTAGCGGACCGCGGACGT 728

Search completed: November 10, 2004, 17:09:33  
Job time : 962.753 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 17:09:58 ; Search time 173.81 Seconds  
(without alignments)  
6187.359 Million cell updates/sec

Title: US-09-909-207-10  
Perfect score: 1513  
Sequence: 1 AAATGTAATTGTATATCT.....TGAAACACCTCCGTCACCTAG 1513

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/prodata/1/ina/5A-COMB.seq: \*  
2: /cgn2\_6/prodata/1/ina/5B-COMB.seq: \*  
3: /cgn2\_6/prodata/1/ina/6A-COMB.seq: \*  
4: /cgn2\_6/prodata/1/ina/6B-COMB.seq: \*  
5: /cgn2\_6/prodata/1/ina/PCTUS-COMB.seq: \*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1513	100.0	1513	3	US-08-470-953A-10
2	1513	100.0	1513	3	US-08-470-953A-11
3	840.6	55.6	871	1	US-08-698-978-1
4	744	49.2	744	3	US-08-470-953A-4
5	744	49.2	744	3	US-08-470-953A-5
6	713.6	47.2	744	3	US-09-189-060B-1
7	673.6	44.5	744	3	US-08-501-126-18
8	663	43.8	663	3	US-08-470-953A-1
9	663	43.8	663	3	US-08-470-953A-2
10	619	40.9	619	3	US-08-470-953A-12
11	587.2	38.8	744	3	US-09-189-060B-11
12	282.8	18.7	1022	3	US-08-275-526C-1
13	282.8	18.7	1022	3	US-08-275-526C-35
14	282.8	18.7	1022	4	US-09-076-677-1
15	282.8	18.7	1022	4	US-09-076-677-35
16	282.8	18.7	1022	4	US-09-073-055-1
17	282.8	18.7	1022	4	US-09-073-055-35
18	272.8	18.0	681	3	US-08-275-526C-30
19	272.8	18.0	681	3	US-08-275-526C-32
20	272.8	18.0	681	4	US-09-076-677-30
21	272.8	18.0	681	4	US-09-076-677-32
22	272.8	18.0	681	4	US-09-073-055-30
23	272.8	18.0	681	4	US-09-073-055-32
24	255.6	16.9	2364	3	US-09-390-234-23
25	255.6	16.9	2364	4	US-09-603-311-23
26	255	16.9	600	3	US-08-275-526C-26
27	255	16.9	600	3	US-08-275-526C-34

Sequence 26, Appl  
Sequence 34, Appl  
Sequence 26, Appl  
Sequence 34, Appl  
Sequence 12, Appl  
Sequence 15, Appl  
Sequence 16, Appl  
Sequence 13, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 53, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 2, Appl

28 255 16.9 600 4 US-09-076-677-26  
29 255 16.9 600 4 US-09-076-677-34  
30 255 16.9 600 4 US-09-073-055-26  
31 255 16.9 600 4 US-09-073-055-34  
32 204.8 13.5 1244 3 US-08-591-685-12  
33 154.4 10.2 164 3 US-08-501-126-15  
34 154.4 10.2 164 3 US-08-501-126-16  
35 150 9.9 150 3 US-08-470-953A-13  
36 132.4 8.8 294 3 US-08-817-946-4  
37 132.4 8.8 294 4 US-09-639-354A-4  
38 119.4 7.9 1375 2 US-08-458-812-1  
39 119.4 7.9 1375 3 US-08-590-563-1  
40 119.4 7.9 1375 4 US-09-770-621-1  
41 119.4 7.9 1375 4 US-09-235-832-1  
42 119 7.9 229 3 US-09-189-060B-53  
43 118.4 7.8 573 1 US-08-709-912-18  
44 118.4 7.8 573 2 US-09-047-370-18  
45 117.8 7.8 1207 1 US-08-575-964-2

#### ALIGNMENTS

RESULT 1  
US-08-470-953A-10  
; Sequence 10, Application US/08470953A  
; Patent No. 6346407

; GENERAL INFORMATION:

; APPLICANT: ANDREE LAHAYE

; APPLICANT: ERIC DE BUYL

; APPLICANT: PIERRE LEDOUX

; APPLICANT: RENE DETROZ

; TITLE OF INVENTION: Xylanase, microorganisms produced it,

; TITLE OF INVENTION: DNA molecule, processes for preparation of this xylanase

; TITLE OF INVENTION: and uses thereof

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE

; STREET: 2000 K St., N.W., Suite 200

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20006

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC Compatible

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/470,953A

; FILING DATE: 6-OCTOBER-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Wilhem F. Gadiano, Esq.

; REGISTRATION NUMBER: 37,136

; REFERENCE/DOCKET NUMBER: 4121-40

; TELEPHONE: 202-429-0625

; TELEFAX: (202) 293-1850

; TELEX: 650 383-5605

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1513 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; ORIGINAL SOURCE:

; STRAIN: Bacillus

; US-08-470-953A-10

Qy	1021	AGAGCAACGCCCTAAGGGGACCAATCACTGTTGATGAGGAAACATATGATCTTACAGAC	1080
Db	1021	AGGAGCAACGCCCTAAGGGGACCAATCACTGTTGATGAGGAAACATATGATCTTACAGAC	1080
Qy	1081	TCCTTAGAGTCAATCAACCCCTCCAAATTAAGGGGATTCGCCACATTTTAAACAATATTTGGAAGTGT	1140
Db	1081	TCCTTAGAGTCAATCAACCCCTCCAAATTAAGGGGATTCGCCACATTTTAAACAATATTTGGAAGTGT	1140
Qy	1141	TCGAAGATCGAAACGACGAGTGGCACGATTTCTGTGACGCAACCACTTTTAGAGCGTGGGA	1200
Db	1141	TCGAAGATCGAAACGACGAGTGGCACGATTTCTGTGACGCAACCACTTTTAGAGCGTGGGA	1200
Qy	1201	AAACTTAGGATGAATATATGGGGAAAAATGATAGAGTCGCGCTTACTGTGTAAGAGGCTATCA	1260
Db	1201	AAACTTAGGATGAATATATGGGGAAAAATGATAGAGTCGCGCTTACTGTGTAAGAGGCTATCA	1260
Qy	1261	AGCTAGCGGAAGTGCTAAATGTATATAGCAATACACTAAGAAATTAACCGGTAACCCCTCTCTC	1320
Db	1261	AGCTAGCGGAAGTGCTAAATGTATATAGCAATACACTAAGAAATTAACCGGTAACCCCTCTCTC	1320
Qy	1321	AACTATTAGTAATGACGAGAGCAATACTTTGGATAAAAAACAATTTAAATATCTTATCTCT	1380
Db	1321	AACTATTAGTAATGACGAGAGCAATACTTTGGATAAAAAACAATTTAAATATCTTATCTCT	1380
Qy	1381	TTCCGTTAGTCTCAATATTTTCAAATAACCTCCCGTTGGATCTTTTCCAAACGGGAGG	1440
Db	1381	TTCCGTTAGTCTCAATATTTTCAAATAACCTCCCGTTGGATCTTTTCCAAACGGGAGG	1440
Qy	1441	TTTTATTGGAAGGTTAAGTATAGTATCTCCGATTCATCCAGAGGAATGCTTTGAAACA	1500
Db	1441	TTTTATTGGAAGGTTAAGTATAGTATCTCCGATTCATCCAGAGGAATGCTTTGAAACA	1500
Qy	1501	CCTCCGTCACCTAG 1513	
Db	1501	CCTCCGTCACCTAG 1513	

RESULT 2  
 US-08-470-953A-11  
 ; Sequence 11, Application US/08470953A  
 ; Patent No. 6346407  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ANDREE LAHAYE  
 ; APPLICANT: ERIC DE BUYL  
 ; APPLICANT: PIERRE LEDOUX  
 ; APPLICANT: RENE DETROZ  
 ; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
 ; TITLE OF INVENTION: DNA molecule, processes for preparation of this xylan  
 ; TITLE OF INVENTION: and uses thereof  
 ; NUMBER OF SEQUENCES: 29  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
 ; STREET: 2000 K St., N.W., Suite 200  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20006  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/470,953A  
 ; FILING DATE: 6-OCTOBER-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Wilhem F. Gadiano, Esq.  
 ; REGISTRATION NUMBER: 37,136  
 ; REFERENCE/DOCKET NUMBER: 4121-40

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1513 base pairs  
TYPE: nucleic acid  
STRAINEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: Bacillus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 620..1363  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 701..1363  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 620..700  
US-08-470-953A-11

Query Match 100.0%; Score 1513; DB 3; Length 1513;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AAATTGAATTTGTATATCTAATGATAACGACAAATCGTCACCTGTTTAAATCTTC 60  
DB 1 AAATTGAATTTGTATATCTAATGATAACGACAAATCGTCACCTGTTTAAATCTTC 60  
QY 61 AAACCAATCTCTTTATTTAAGCGTAACGACCTTGCAATCTTATACAGAAACATCTTT 120  
DB 61 AAACCAATCTCTTTATTTAAGCGTAACGACCTTGCAATCTTATACAGAAACATCTTT 120  
QY 121 ATAGGAATCTCCATTTGCAAGACGATGAAATCTTTTCCCTATTTTATCTTATCG 180  
DB 121 ATAGGAATCTCCATTTGCAAGACGATGAAATCTTTTCCCTATTTTATCTTATCG 180  
QY 181 CCTTGATCGGTTTAAATTTGTAAATCTTTATTTAGTTTACGTGATGTTCCCTCATTCAT 240  
DB 181 CCTTGATCGGTTTAAATTTGTAAATCTTTATTTAGTTTACGTGATGTTCCCTCATTCAT 240  
QY 241 CATTAATCAAGTTAAGCGTAGAGTCACTTTTTCGGTTCCTCAAAATACCTGAAGAAC 300  
DB 241 CATTAATCAAGTTAAGCGTAGAGTCACTTTTTCGGTTCCTCAAAATACCTGAAGAAC 300  
QY 301 ATTTATGTCATATTTCTCAGCGCGCTCCATATGGAATATATATATCTTTTATACATA 360  
DB 301 ATTTATGTCATATTTCTCAGCGCGCTCCATATGGAATATATATATCTTTTATACATA 360  
QY 361 TTAAGTAAATAGTATATATCTTGGTTTATCAAAATGTGAGATAATCTAATTTGATCAACA 420  
DB 361 TTAAGTAAATAGTATATATCTTGGTTTATCAAAATGTGAGATAATCTAATTTGATCAACA 420  
QY 421 AGCAGCTATCCAAAAACACATGATGTTGACCTCTTAAAGAAAGTGCTATCTATGAAA 480  
DB 421 AGCAGCTATCCAAAAACACATGATGTTGACCTCTTAAAGAAAGTGCTATCTATGAAA 480  
QY 481 GATAATTTCCAGTTTCAAAATTTGAATAGTGTGATGGAATAGTTTGAATGTCAACTG 540  
DB 481 GATAATTTCCAGTTTCAAAATTTGAATAGTGTGATGGAATAGTTTGAATGTCAACTG 540  
QY 541 CTGTGAAGAGGGGTAGTAGTACCGTAGACTTCAATACCAAAATAGTTTGAATAAAA 600  
DB 541 CTGTGAAGAGGGGTAGTAGTACCGTAGACTTCAATACCAAAATAGTTTGAATAAAA 600  
QY 601 TTAAGAGGGAATGCGCTAATGAGACAAAAGAAATGACGTGATTTAGCTCTTTTAGT 660  
DB 601 TTAAGAGGGAATGCGCTAATGAGACAAAAGAAATGACGTGATTTAGCTCTTTTAGT 660  
QY 661 TTGTTTTGACTAACCTTACCTGACAGAAATTAATTCAGGCACAAATCGTCACGCAATTC 720

DB 661 TTGTTTTGACTAACCTTACCTGACAGAAATAATTCAGGCACAAATCGTCACGCAATTC 720  
QY 721 CATTTGGCAACACGATGCTATGATTAATTTTGGAAAGATAGCGGTGGCTCTTGGGAC 780  
DB 721 CATTTGGCAACACGATGCTATGATTAATTTTGGAAAGATAGCGGTGGCTCTTGGGAC 780  
QY 781 AATGATTTCTCAATCATGCGGTAGCTTCAGTGCCTCAATGGAACAAATGTTAAACAATATT 840  
DB 781 AATGATTTCTCAATCATGCGGTAGCTTCAGTGCCTCAATGGAACAAATGTTAAACAATATT 840  
QY 841 ATTTCGTTAAAGGTAAATAATTTCAATGAAAACACACCAACCAAGTTGGTAAACATGTC 900  
DB 841 ATTTCGTTAAAGGTAAATAATTTCAATGAAAACACACCAACCAAGTTGGTAAACATGTC 900  
QY 901 CATAAACTACGAGCGCAACTTCCAAACCAATGTTATGCTATTTATGCTCTATGGTTG 960  
DB 901 CATAAACTACGAGCGCAACTTCCAAACCAATGTTATGCTATTTATGCTCTATGGTTG 960  
QY 961 GACTGTTGACCTCTTGTGCAATATTATTTGTCACAGTTGGGCAACTGGCGTCCACC 1020  
DB 961 GACTGTTGACCTCTTGTGCAATATTATTTGTCACAGTTGGGCAACTGGCGTCCACC 1020  
QY 1021 AGGACCAACGCTAAAGGGACCATCACTGTTGATGAGGAAACATATGATCTACGAGAC 1080  
DB 1021 AGGACCAACGCTAAAGGGACCATCACTGTTGATGAGGAAACATATGATCTACGAGAC 1080  
QY 1081 TCTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCAATTTAAACAATATTGAGTGT 1140  
DB 1081 TCTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCAATTTAAACAATATTGAGTGT 1140  
QY 1141 TCGAAGATCGAAACGACGAGTGGCAAGATTTCTGTCAGCAACCACTTTAGAGCGTGGGA 1200  
DB 1141 TCGAAGATCGAAACGACGAGTGGCAAGATTTCTGTCAGCAACCACTTTAGAGCGTGGGA 1200  
QY 1201 AAATTTAGGATGAATATGAGGAAATATGATGAGTTCGCGCTTACTGTAGAAGGCTATCA 1260  
DB 1201 AAATTTAGGATGAATATGAGGAAATATGATGAGTTCGCGCTTACTGTAGAAGGCTATCA 1260  
QY 1261 AAGTACGGAAGTGTAAATGATATAGCAATACACCTAAGAAATTAACGGTAACCTCTCTC 1320  
DB 1261 AAGTACGGAAGTGTAAATGATATAGCAATACACCTAAGAAATTAACGGTAACCTCTCTC 1320  
QY 1321 AACTATTAGTAATGACGAGAGCATAACTTTGGATAAAAACAATTAATAATCTTATCTCT 1380  
DB 1321 AACTATTAGTAATGACGAGAGCATAACTTTGGATAAAAACAATTAATAATCTTATCTCT 1380  
QY 1381 TTCCGTTTCACTCTCAATTTTCAATAAACCCTCCCGGTGGATCTTTTCCACGCGGAGG 1440  
DB 1381 TTCCGTTTCACTCTCAATTTTCAATAAACCCTCCCGGTGGATCTTTTCCACGCGGAGG 1440  
QY 1441 TTTTATTGGAAGGTTAAGTATAGTATCTCCGATTTCCATCCAGAGGAATGCTTGAACA 1500  
DB 1441 TTTTATTGGAAGGTTAAGTATAGTATCTCCGATTTCCATCCAGAGGAATGCTTGAACA 1500  
QY 1501 CCTCCGCTACTAG 1513  
DB 1501 CCTCCGCTACTAG 1513

RESULT 3

US-08-698-978-1  
; Sequence 1, Application US/08698978  
; Patent No. 5770424  
; GENERAL INFORMATION:  
; APPLICANT: Outtrup, Helle  
; APPLICANT: Dammann, Claus  
; APPLICANT: Olsen, Arne Agerlin  
; APPLICANT: Bisg rd-Frantzen, Henrik  
; APPLICANT: Sch lein, Martin  
; APPLICANT: Jorgensen, Per Linea  
; TITLE OF INVENTION: DNA CONSTRUCTS AND METHODS OF PRODUCING  
; ; XYLANOLYTIC ENZYMES

; NUMBER OF SEQUENCES: 2  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: No. 5770424th No. 5770424th disk of No. 5770424th America, Inc.  
 ; STREET: 405 Lexington Avenue, Suite 6400  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10174-6401  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/698,978  
 ; FILING DATE: 16 August 1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/470,398  
 ; FILING DATE: 06-JUN-1996  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/343,600  
 ; FILING DATE: 30-NOV-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/DK93/00218  
 ; FILING DATE: 2-JUL-92  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Gregg, Valeta A.  
 ; REGISTRATION NUMBER: 35,127  
 ; REFERENCE/DOCKET NUMBER: 3794.400-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212 867 0123  
 ; TELEFAX: 212 867 0298  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 871 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Bacillus ageraderens  
 ; STRAIN: NCIMB 40482  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 82..744  
 ; US-08-698-978-1  
 ;  
 ; Query Match 55.6%; Score 840.6; DB 1; Length 871;  
 ; Best Local Similarity 97.8%; Pred. No. 5.8e-216;  
 ; Matches 852; Conservative 0; Mismatches 19; Indels 0; Gaps 0;  
 ;  
 QY 620 ATGACACAAAGAAATGACGTTGATTTTGTAGCTTTTGTAGTTTGTGCTACTAACCTTA 679  
 Db 1 ATGACACAAAGAAATGACGTTGATTTTGTAGCTTTTGTAGTTTGTGCTACTAACCTTA 60  
 ;  
 QY 680 CTTGACAGAAATATTCAGGCAAAATCTGACGCAAAATTCATTCGACCAATCCAGATGC 739  
 Db 61 CTTGACAGAAATATTCAGGCAAAATCTGACGCAAAATTCATTCGACCAATCCAGATGC 120  
 ;  
 QY 740 TATGATTATGATTTTGGAAAGATAGCGTGTGCTCTGGGACAAATGATTTCTCAATCATGCC 799  
 Db 121 TATGATTATGATTTTGGAAAGATAGCGTGTGCTCTGGGACAAATGATTTCTCAATCATGCC 180  
 ;  
 QY 800 GGTAGTTTCACTGCTCCCAATGGAACAAATGTTTAAACATATTTATTCGTAAGGTAAGAAA 859  
 Db 181 GGTAGTTTCACTGCTCCCAATGGAACAAATGTTTAAACATATTTATTCGTAAGGTAAGAAA 240  
 ;  
 QY 860 TTCAATGAAACACAAACACCAACCAAGTTGTGTAACATGTCATTAATCTAGCGAGCCAAC 919  
 Db 241 TTCATGAAACACAAACACCAACCAAGTTGTGTAACATGTCATTAATCTAGCGAGCCAAC 300  
 ;  
 QY 920 TTCCAAACAAATGGTAAATGCGTATTTATGCGTCTATGTTGACTGTTGACCTCTTGTGTC 979

Db 301 TTCCAGCAACGGAATGCGTATTTATGCGTCTATGTTGACTGTTGACCTCTTGTGTC 360  
 QY 980 GAATATTATTTGTCGACAGTTGGGGCAACTGGGCTCCACAGGAGCAACGCTTAAGGGG 1039  
 Db 361 GAATATTATTTGTCGACAGTTGGGGCAACTGGGCTCCACAGGAGCAACGCTTAAGGGA 420  
 QY 1040 ACCATCACTGTTGATGGAGGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 1099  
 Db 421 ACCATCACTGTTGATGGAGGAACATATGATATCTATGAAACTCTTAGAGTCAATCAGCCC 480  
 QY 1100 TCCATTAAAGGGATTGCCACATTTTAAACAATTTGGAGTGTTCGAAGATCGAAACGCGAG 1159  
 Db 481 TCCATTAAAGGGATTGCCACATTTTAAACAATTTGGAGTGTTCGAAGATCGAAACGCGAG 540  
 QY 1160 AGTGCGACGATTTCTGTGACCAACCACTTTTAGAGCGTGGGAAACCTTAGGAGTGAATATG 1219  
 Db 541 AGTGCGACCAATTTCTGTGACCAACCACTTTTAGAGCGTGGGAAACCTTAGGAGTGAATATG 600  
 QY 1220 GGGAAATGATGAAGTCCGCTTACTGTAGAAGGCTATCAAGTAGCGGAAGTCTTAAT 1279  
 Db 601 GGGAAATGATGAAGTCCGCTTACTGTAGAAGGCTATCAAGTAGCGGAAGTCTTAAT 660  
 QY 1280 GTATATAGCAATACACTAAGAAATTAACGCTTAACCTCTCAACTATTAGTAATGACGAG 1339  
 Db 661 GTATATAGCAATACACTAAGAAATTAACGCTTAACCTCTCAACTATTAGTAATGACGAG 720  
 QY 1340 AGCATAACTTTGGATAAAACAATTAATAATCTCTTTCGTTTCAGTCTCTCAATTA 1399  
 Db 721 AGCATAACTTTAGATAAAACAATTAATAATCTCTTTCGTTTCAGTCTCTCAATTA 780  
 QY 1400 TTTTCAATAACCTCCCGGTTGGATCTTTTCAACGGGAGGTTTATTTGGAAGGTTAAG 1459  
 Db 781 TTTTCAATAACCTCCCGGTTGGATCTTTTCAACGGGAGGTTTATTTGGAAGGTTAAG 840  
 QY 1460 TATAGTATACCTCCGATTCCTCCATCCAGAGGAAT 1490  
 Db 841 TATAGTATACCTCCGATTCCTCCATCCAGAGGAAT 871  
 ;  
 RESULT 4  
 ; US-08-470-953A-4  
 ; Sequence 4, Application US/08470953A  
 ; Patent No. 6346407  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ANDREE LAHAYE  
 ; APPLICANT: ERIC DE BUYL  
 ; APPLICANT: PIERRE LEDOUX  
 ; APPLICANT: RENE DETROZ  
 ; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
 ; TITLE OF INVENTION: DNA molecule, processes for preparation of this xylanase  
 ; NUMBER OF INVENTION: and uses thereof  
 ; NUMBER OF SEQUENCES: 29  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: WILLIAM BRINKS HOPFER GILSON & LIONE  
 ; STREET: 2000 K St., N.W., Suite 200  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20006  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/470,953A  
 ; FILING DATE: 6-OCTOBER-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:



NAME: Wilhem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-40  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 744 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: Bacillus  
US-08-470-953A-4

Query Match 49.2%; Score 744; DB 3; Length 744;  
Best Local Similarity 100.0%; Pred. No. 4.6e-190; Indels 0; Gaps 0;  
Matches 744; Conservative 0; Mismatches 0;

```
QY 620 ATGAGCAAAAGAAATTGACGTTGATTTTAGCCCTTTTGTAGTTTGTGCTTAACCTTA 679
DB 1 ATGAGCAAAAGAAATTGACGTTGATTTTAGCCCTTTTGTAGTTTGTGCTTAACCTTA 60

QY 680 CTGCGAGAAATAATTGAGGCAAAATCGTCACCGCAAAATTCATTCGTCGCAACGATGGC 739
DB 61 CTGCGAGAAATAATTGAGGCAAAATCGTCACCGCAAAATTCATTCGTCGCAACGATGGC 120

QY 740 TATGATTGAAATTTTGGAAAGATAGCGTGGCTCTGGGACATGATTTCTCAATCATGGC 799
DB 121 TATGATTGAAATTTTGGAAAGATAGCGTGGCTCTGGGACATGATTTCTCAATCATGGC 180

QY 800 GGTACGTTTCAGTGCCCAATGGAACATGTTAAACAATATTTCCGTAAAGTAAAAA 859
DB 181 GGTACGTTTCAGTGCCCAATGGAACATGTTAAACAATATTTCCGTAAAGTAAAAA 240

QY 860 TTCAATGAAACACAAACACACCAAGTTGGTAAACATGTCATTAACCTACGAGGCAAC 919
DB 241 TTCAATGAAACACAAACACACCAAGTTGGTAAACATGTCATTAACCTACGAGGCAAC 300

QY 920 TTCACCAAAATGTAATGCTATTTATGCTGCTATGTTGGTGGTGTGACCTCTTGTG 979
DB 301 TTCACCAAAATGTAATGCTATTTATGCTGCTATGTTGGTGGTGTGACCTCTTGTG 360

QY 980 GAATATTATTTGTCAGACGTTGGGCAACTGGGCTCCACGAGGCAACGCTTAAGGGG 1039
DB 361 GAATATTATTTGTCAGACGTTGGGCAACTGGGCTCCACGAGGCAACGCTTAAGGGG 420

QY 1040 ACCATCACTGTTGATGAGGAAACATATGATATCTACGAGCTCTTAGAGTCAATCAACC 1099
DB 421 ACCATCACTGTTGATGAGGAAACATATGATATCTACGAGCTCTTAGAGTCAATCAACC 480

QY 1100 TCCATTAAGGGGATTCGACATTTAAACATATTTGGAGTGTGGAAGTCGAAACGCAACG 1159
DB 481 TCCATTAAGGGGATTCGACATTTAAACATATTTGGAGTGTGGAAGTCGAAACGCAACG 540

QY 1160 AGTGGCACGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATG 1219
DB 541 AGTGGCACGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATG 600

QY 1220 GGGAAATATGATGAAGTCGGCTTATCTGTAGAGGCTATCAAGTAGCGGAAGTCTAAT 1279
DB 601 GGGAAATATGATGAAGTCGGCTTATCTGTAGAGGCTATCAAGTAGCGGAAGTCTAAT 660

QY 1280 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCACTATTAGTATGACGAG 1339
DB 661 GTATATAGCAATACACTAAGAAATTAACGGTAAACCTCTCTCACTATTAGTATGACGAG 720

QY 1340 AGCATAACTTTGGATAAAAAACAAT 1363
DB 721 AGCATAACTTTGGATAAAAAACAAT 744
```

## RESULT 5

US-08-470-953A-5  
Sequence 5, Application US/08470953A  
Patent No. 6346407

## GENERAL INFORMATION:

APPLICANT: ANDREE LAHAYE  
APPLICANT: ERIC DE BUYL  
APPLICANT: PIERRE LEDOUX  
APPLICANT: RENE DETROZ

TITLE OF INVENTION: Xylanase, microorganisms produced it,  
TITLE OF INVENTION: DNA molecule, processes for preparation of this xylanase  
TITLE OF INVENTION: and uses thereof  
NUMBER OF SEQUENCES: 29

## CORRESPONDENCE ADDRESS:

ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
STREET: 2000 K St., N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/470,953A  
FILING DATE: 6-OCTOBER-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

## FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Wilhem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-40  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 744 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: Bacillus

## FEATURE:

NAME/KEY: CDS  
LOCATION: 1..744  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 82..744  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 1..81

US-08-470-953A-5

Query Match 49.2%; Score 744; DB 3; Length 744;

Best Local Similarity 100.0%; Pred. No. 4.6e-190; Indels 0; Gaps 0;

Matches 744; Conservative 0; Mismatches 0;

```
QY 620 ATGAGCAAAAGAAATTGACGTTGATTTTAGCCCTTTTGTAGTTTGTGCTTAACCTTA 679
DB 1 ATGAGCAAAAGAAATTGACGTTGATTTTAGCCCTTTTGTAGTTTGTGCTTAACCTTA 60

QY 680 CTGCGAGAAATAATTGAGGCAAAATCGTCACCGCAAAATTCATTCGTCGCAACGATGGC 739
DB 61 CTGCGAGAAATAATTGAGGCAAAATCGTCACCGCAAAATTCATTCGTCGCAACGATGGC 120
```





STRAIN: Bacillus  
US-08-470-953A-1

Query Match 43.8%; Score 663; DB 3; Length 663;  
Best Local Similarity 100.0%; Pred. No. 2.4e-168;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 701 CAAATCGTCACGACAAATCCATTTGGCAACACGATGGTATGATTAATGAATTTGGAAA 760  
DB 1 CAAATCGTCACGACAAATCCATTTGGCAACACGATGGTATGATTAATGAATTTGGAAA 60

QY 761 GATAGCGGTGGTCTCGGACAAATGATTTCTCAATCATGCGCGGTACGTTCACTGCGCCCAATGG 820  
DB 61 GATAGCGGTGGTCTCGGACAAATGATTTCTCAATCATGCGCGGTACGTTCACTGCGCCCAATGG 120

QY 821 AACAAATGTTAAACAAATATTTTCGTAAGGTAAATAATTAATCAATGAACACACACAC 880  
DB 121 AACAAATGTTAAACAAATATTTTCGTAAGGTAAATAATTAATCAATGAACACACACAC 180

QY 881 CAACAAATGTTAAACAAATGTTTCGTAAGGTAAATAATTAATCAATGAACACACACAC 940  
DB 181 CAACAAATGTTAAACAAATGTTTCGTAAGGTAAATAATTAATCAATGAACACACACAC 240

QY 941 TATTTATGCGTCTATGTTGGACTGTTGACCCCTTTGTCGAATATTAATTTGTCGACAGT 1000  
DB 241 TATTTATGCGTCTATGTTGGACTGTTGACCCCTTTGTCGAATATTAATTTGTCGACAGT 300

QY 1001 TGGGGCAACTGGCGTCCACGAGGACCACTTCCAAACCAATGTAATGG 1060  
DB 301 TGGGGCAACTGGCGTCCACGAGGACCACTTCCAAACCAATGTAATGG 360

QY 1061 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTCACCA 1120  
DB 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTCACCA 420

QY 1121 TTTAAACAATATTGGAGTGTTCGAAGATCGAAGACGACGAGTGGCAGCATTTCTGTGAGC 1180  
DB 661 AAT 663

RESULT 9  
US-08-470-953A-2  
Sequence 2, Application US/08470953A  
Patent No. 6346407

GENERAL INFORMATION:  
APPLICANT: ANDREE LAHAYE  
APPLICANT: ERIC DE BUIL  
APPLICANT: PIERRE LEDOUX  
APPLICANT: RENE DETROZ

TITLE OF INVENTION: Xylanase, microorganisms produced it,  
TITLE OF INVENTION: DNA molecule, processes for preparation of this xylanase  
TITLE OF INVENTION: and uses thereof  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
STREET: 2000 K St., N.W., Suite 200  
CITY: Washington

STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/470,953A  
FILING DATE: 6-OCTOBER-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Wilhem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-40  
TELEPHONE: 202-429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 663 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: Bacillus  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 1..663  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..663  
US-08-470-953A-2

Query Match 43.8%; Score 663; DB 3; Length 663;  
Best Local Similarity 100.0%; Pred. No. 2.4e-168;  
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 701 CAAATCGTCACGACAAATCCATTTGGCAACACGATGGTATGATTAATGAATTTGGAAA 760  
DB 1 CAAATCGTCACGACAAATCCATTTGGCAACACGATGGTATGATTAATGAATTTGGAAA 60

QY 761 GATAGCGGTGGTCTCGGACAAATGATTTCTCAATCATGCGCGGTACGTTCACTGCGCCCAATGG 820  
DB 61 GATAGCGGTGGTCTCGGACAAATGATTTCTCAATCATGCGCGGTACGTTCACTGCGCCCAATGG 120

QY 821 AACAAATGTTAAACAAATATTTTCGTAAGGTAAATAATTAATCAATGAACACACACAC 880  
DB 121 AACAAATGTTAAACAAATATTTTCGTAAGGTAAATAATTAATCAATGAACACACACAC 180

QY 881 CAACAAATGTTAAACAAATGTTTCGTAAGGTAAATAATTAATCAATGAACACACACAC 940  
DB 181 CAACAAATGTTAAACAAATGTTTCGTAAGGTAAATAATTAATCAATGAACACACACAC 240

QY 941 TATTTATGCGTCTATGTTGGACTGTTGACCCCTTTGTCGAATATTAATTTGTCGACAGT 1000  
DB 241 TATTTATGCGTCTATGTTGGACTGTTGACCCCTTTGTCGAATATTAATTTGTCGACAGT 300

QY 1001 TGGGGCAACTGGCGTCCACGAGGACCACTTCCAAACCAATGTAATGG 1060  
DB 301 TGGGGCAACTGGCGTCCACGAGGACCACTTCCAAACCAATGTAATGG 360

QY 1061 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTCACCA 1120  
DB 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTCACCA 420

QY 1121 TTTAAACAATATTGGAGTGTTCGAAGATCGAAGACGACGAGTGGCAGCATTTCTGTGAGC 1180

Db 421 TTTAAACATATTTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAACGATTTCTGTGACG 480  
QY 1181 AACCACTTTAGAGCGTGGGAAACCTTAGGATGAATATGGGAAATGTATGAAGTCGG 1240  
Db 481 AACCACTTTAGAGCGTGGGAAACCTTAGGATGAATATGGGAAATGTATGAAGTCGG 540  
QY 1241 CTTACTGTAGAAGCTTCAAAAGTAGCGGAAGTCTAAATGTATATAGCAATACACTAAGA 1300  
Db 541 CTTACTGTAGAAGCTTCAAAAGTAGCGGAAGTCTAAATGTATATAGCAATACACTAAGA 600  
QY 1301 ATTACGGTAACCCCTCTCTCAACTATTAGTAATGACGAGACATAACTTTGGATAAAAA 1360  
Db 601 ATTACGGTAACCCCTCTCTCAACTATTAGTAATGACGAGACATAACTTTGGATAAAAA 660  
QY 1361 AAT 1363  
Db 661 AAT 663  
RESULT 10  
US-08-470-953A-12  
; Sequence 12, Application US/08470953A  
; Patent No. 6346407  
; GENERAL INFORMATION:  
; APPLICANT: ANDREE LAHAYE  
; APPLICANT: ERIC DE BUYL  
; APPLICANT: PIERRE LEDOUX  
; APPLICANT: RENE DETROZ  
; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
; TITLE OF INVENTION: DNA molecule, processes for preparation of this xylanase  
; TITLE OF INVENTION: and uses thereof  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
; STREET: 2000 K St., N.W., Suite 200  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/470,953A  
; FILING DATE: 6-OCTOBER-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wilhem F. Gadiano, Esq.  
; REGISTRATION NUMBER: 37,136  
; REFERENCE/DOCKET NUMBER: 4121-40  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-429-0625  
; TELEFAX: (202) 293-1850  
; TELEX: 650 383-5605  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 619 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-470-953A-12

Query Match 40.9%; Score 619; DB 3; Length 619;  
Best Local Similarity 100.0%; Pred. No. 1.5e-156;  
Matches 619; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATGTAATGTGTATATCTAATGATAACGACAAATCGTCACTGTGTTTTAAACTAATCTC 60  
Db 1 AAATGTAATGTGTATATCTAATGATAACGACAAATCGTCACTGTGTTTTAAACTAATCTC 60  
QY 61 AAACCAATACCTCTTTTAAACGCTAACCACTTGCATCTTATCACAAGACATCTCTTT 120  
Db 61 AAACCAATACCTCTTTTAAACGCTAACCACTTGCATCTTATCACAAGACATCTCTTT 120  
QY 121 ATAGGAACCTTTCCCAATTTGCAAGACGATAAAAAATCTTTTCCCTATTTTATCTTATCG 180  
Db 121 ATAGGAACCTTTCCCAATTTGCAAGACGATAAAAAATCTTTTCCCTATTTTATCTTATCG 180  
QY 181 CTTTCATCGGTTTAAATTTGTAACCTTTTATTTAGTTTACGTCATGTTCCCTCATTCATAC 240  
Db 181 CTTTCATCGGTTTAAATTTGTAACCTTTTATTTAGTTTACGTCATGTTCCCTCATTCATAC 240  
QY 241 CATTAATCACAGTTAAACGCTAGAGTCATCTTTTTCGGTTCCTCAAAAAATACCTCGAAGAAC 300  
Db 241 CATTAATCACAGTTAAACGCTAGAGTCATCTTTTTCGGTTCCTCAAAAAATACCTCGAAGAAC 300  
QY 301 ATTTATGTCATATTTTCTCACGCCCTCCATAATGGAATATATATATCTTTTATACATA 360  
Db 301 ATTTATGTCATATTTTCTCACGCCCTCCATAATGGAATATATATATCTTTTATACATA 360  
QY 361 TTAAGTAAATTTAGTATATATCTTTCGTTATCAAAATGTGAGATAATCTAATGATCAAAACA 420  
Db 361 TTAAGTAAATTTAGTATATATCTTTCGTTATCAAAATGTGAGATAATCTAATGATCAAAACA 420  
QY 421 AGCAGCTATCCAAAAACACTGATGTTGACCTTTAAAGAGTGCACACTATCTATGAAAAA 480  
Db 421 AGCAGCTATCCAAAAACACTGATGTTGACCTTTAAAGAGTGCACACTATCTATGAAAAA 480  
QY 481 GATAATTTCCAGTTTCAAAATTTGAAATAGTGTGTATGGAATAGTTTGAATGTCAACTG 540  
Db 481 GATAATTTCCAGTTTCAAAATTTGAAATAGTGTGTATGGAATAGTTTGAATGTCAACTG 540  
QY 541 CTGTGAAAGGAGGAGGTAGGTAGTACCGTAGATTCATTACCAAAAAATTTAGTTGTAAAAAAA 600  
Db 541 CTGTGAAAGGAGGAGGTAGGTAGTACCGTAGATTCATTACCAAAAAATTTAGTTGTAAAAAAA 600  
QY 601 TTAAAGAGGAGGAGTGCCTA 619  
Db 601 TTAAAGAGGAGGAGTGCCTA 619  
RESULT 11  
US-09-189-060B-11  
; Sequence 11, Application US/09189060B  
; Patent No. 6270968  
; GENERAL INFORMATION:  
; APPLICANT: Dalboge, Henrik  
; APPLICANT: Sandal, Thomas  
; APPLICANT: Kauppinen, Markus  
; APPLICANT: Borge, Diderichsen  
; TITLE OF INVENTION: Method Of Providing No. 6270968el DNA Sequences  
; FILE REFERENCE: 4772.204-US  
; CURRENT APPLICATION NUMBER: US/09/189,060B  
; PRIOR FILING DATE: 1998-11-10  
; PRIOR FILING DATE: 1997-05-12  
; NUMBER OF SEQ ID NOS: 74  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 744  
; TYPE: DNA  
; ORGANISM: Hybrid  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(744)  
US-09-189-060B-11

Query Match 38.8%; Score 587.2; DB 3; Length 744;  
Best Local Similarity 86.8%; Pred. No. 5.6e-148;

Matches 646; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 620 ATGAGACAAAGAAATGAGCTGATTTAGCTTTTATGTTGCTTGTGCTTACCTTA 679  
 Db |||||  
 QY 1 ATGAGACAAAGAAATGAGCTGATTTAGCTTTTATGTTGCTTGTGCTTACCTTA 60  
 Db |||||  
 QY 680 CTGCGAGAAATTAATTCAGGACAAATTCGTCACGCAATTCATTTGGCAACCAAGATGGC 739  
 Db |||||  
 QY 61 CTGCGAGAAATTAATTCAGGACAAATTCGTCACGCAATTCATTTGGCAACCAAGATGGC 120  
 Db |||||  
 QY 740 TATGATTAATTAATTTGGAAAGATGCGGTGCTCTGGGACAAATGATTTCTCAATCATGGC 799  
 Db |||||  
 QY 121 TATGATTAATTAATTTGGAAAGATGCGGTGCTCTGGGACAAATGATTTCTCAATCATGGC 180  
 Db |||||  
 QY 800 GGTAGCTTCAGTGCCCAATGGAACAAATGTTTACACATATATTCGTTAAAGGTTAAAAA 859  
 Db |||||  
 QY 181 GGTAGCTTCAGTGCCCAATGGAACAAATGTTTACACATATATTCGTTAAAGGTTAAAAA 240  
 Db |||||  
 QY 860 TTCAATGAAACAAACACCAACCAAGTTGCTTAAACATGTCATTAATTCATGAGCCCAAC 919  
 Db |||||  
 QY 241 TTCAATGAAACAAACACCAACCAAGTTGCTTAAACATGTCATTAATTCATGAGCCCAAC 300  
 Db |||||  
 QY 920 TTCCACCAAAATGGAATGCGTATTTATGCGTCTATGTTGACCTTGTGACCTTGTGTC 979  
 Db |||||  
 QY 301 TTCCACCAAAATGGAATGCGTATTTATGCGTCTATGTTGACCTTGTGACCTTGTGTC 360  
 Db |||||  
 QY 980 GAATATTATGTCGACAGTTGGGGCAACTGCGCTCCACGAGGACCAACGCTTAAGGG 1039  
 Db |||||  
 QY 361 GAATATTATGTCGACAGTTGGGGCAACTGCGCTCCACGAGGACCAACGCTTAAGGG 420  
 Db |||||  
 QY 1040 ACCATCACTGTTGATGGAGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 1099  
 Db |||||  
 QY 421 ACCATCACTGTTGATGGAGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480  
 Db |||||  
 QY 1100 TCCATTAAGGGATGCGCAATTTAAACAAATTTGAGTGTTCGAAGATGCGAAGCGACG 1159  
 Db |||||  
 QY 481 TCTATTTCAGGCGACCGCCACCTTCAATCAGTACTGCTGATTCGACAGGCAAGCGGACC 540  
 Db |||||  
 QY 1160 AGTGGCAGATTTCTGTCAGCAACCACTTTAGAGGCTGGGAAACCTTAGGGATGAATATG 1219  
 Db |||||  
 QY 541 ACGGCACTGTCACGCGCAACCACTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 Db |||||  
 QY 1220 GGGAAATGATGAAGTCCGCTTACTGTAGAGGCTATCAAGTAGGGAGTCTTAAT 1279  
 Db |||||  
 QY 601 GGTGCAATCAATTCACAGATCCTGTTACTGAGGCTACCAATCTACCGGAAGTCTAAT 660  
 Db |||||  
 QY 1280 GTATATGCAATACACTAAGAAATTAACGGTAAACCTCTCTCACTATTAGTAATGACGAG 1339  
 Db |||||  
 QY 661 GTATATGCAATACACTAAGAAATTAACGGTAAACCTCTCTCACTATTAGTAATGACGAG 720  
 Db |||||  
 QY 1340 AGCATAACTTTGGATAAAAAACAAT 1363  
 Db |||||  
 QY 721 AGCATAACTTTAGATAAAAAACAAT 744  
 Db |||||

RESULT 12  
 US-08-275-526C-1  
 ; Sequence 1, Application US/08275526C  
 ; Patent No. 6180382  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DE BUYL, ERIC  
 ; APPLICANT: LAHAYE, ANDR E  
 ; APPLICANT: LEDOUX, PIERRE  
 ; APPLICANT: AMORY, ANTOINE  
 ; APPLICANT: DETROZ, REN  
 ; APPLICANT: ANDRE, CHRISTOPHE  
 ; APPLICANT: VETTER, ROMAN  
 ; TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,  
 ; TITLE OF INVENTION: EXPRESSION VECTORS FOR SUCH XYLANASE AND  
 ; TITLE OF INVENTION: OTHER PROTEINS, HOST ORGANISMS THEREFOR AND  
 ; TITLE OF INVENTION: USE THEREOF  
 ; NUMBER OF SEQUENCES: 35  
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.  
 STREET: 2000 K St., N.W., Suite 200  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20006  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/275,526C  
 FILING DATE: 15-JUL-1994  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Gadiano, Wilhem F.  
 REGISTRATION NUMBER: 37,136  
 REFERENCE/DOCKET NUMBER: 4121-49  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 429-0625  
 TELEFAX: (202) 293-0625  
 TELEX: 650 383 5605  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1022 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: genomic DNA  
 ORIGINAL SOURCE:  
 ORGANISM: Bacillus pumilus  
 STRAIN: PRL B12  
 US-08-275-526C-1

Query Match 18.7%; Score 282.8; DB 3; Length 1022;  
 Best Local Similarity 62.1%; Pred. No. 3.4e-66;  
 Matches 481; Conservative 0; Mismatches 287; Indels 6; Gaps 2;

QY 599 AATTAAAGGAGGAATGCCTAATGAGACAAAGAAATGAGCTTGTGATTTAGCTTTTAA 658  
 Db |||||  
 QY 165 AAAAAGAGGAGGATGGAATGAATTTGAAAGATTTAGGCTGTTGTTGTGATGTGT 224  
 Db |||||  
 QY 659 GTTTGTTTTCACATAACCTTACCTGACAAATAATTCAGGCACAAATCGTCACCGACAAT 718  
 Db |||||  
 QY 225 ATTGATTTGCTGACACTGACGCTGTCGGCTCATGCGGAAACGATTTATGATAT 284  
 Db |||||  
 QY 719 TCCATTGCGCAACCAAGTGTGCTATGATTTGAAATTTTGGAAAGATAGCGGTGCTCTGG 778  
 Db |||||  
 QY 285 AGGATAGGACACACAGCGGATACGATTTTGAATTTATGGAAGGATTAC---GGAAATACC 341  
 Db |||||  
 QY 779 ACATGATTTCAATCATGCGGTGACGTTAGTCCCAATGGAACATGTTAAACACATA 838  
 Db |||||  
 QY 342 TCGATGACACTCAATAACGCGGGCATTTAGTCAACAGCTGGAACAAATTTGGAATGCC 401  
 Db |||||  
 QY 839 TTATTCGTAAGGTAAAAAATTCATGAAACACAAACACCAACCAAGTTGGTAACATG 898  
 Db |||||  
 QY 402 TTATTCGAAAGGAAAGAGTTGATTCACCTAAACTCATCATCACTTTGGCAACATC 461  
 Db |||||  
 QY 899 TCCATAAACTACGAGGCAACTTCCAAACCAATTTGTAATGCTATTTATGCGTCTATGGT 958  
 Db |||||  
 QY 462 TCCATCAACTACACACGAGCTTTAAACCGCGGGGAAATTCCTATTATGTCGTATGGC 521  
 Db |||||  
 QY 959 TGGACTGTGACCTCTTGTGCAATATATATTTGTCCACAGTTGGGGCAACTGCGCTCCA 1018  
 Db |||||  
 QY 522 TGGACAAATCTCCATTAGCTGAATACATATTGTTAGTCACTGAGGCGACATATCGTCCA 581  
 Db |||||  
 QY 1019 CCAGGACCAACGCTAAGGGGACCATCACTGTTGATGAGGAAACATATGATATCTACGAG 1078  
 Db |||||  
 QY 582 ACAG---ACGTATAAGGATCATTTTATGCCGATGAGGACATATGACATATATGAA 638  
 Db |||||  
 QY 1079 ACTCTTAGAGTCAATCAACCTCCATTAAGGGATTCGCCACATTTAAACAAATATTGAGT 1138  
 Db |||||

Db 639 ACGCTCCGTGCTCAATCAGCCTTCTATCATTTGGAGAGCTACCTTCAAAACAATTTGGAGT 698  
 QY 1139 GTTCGAAGATCGAAGCGACGAGTGTGTCAGCAACCACTTTAGAGCGTGG 1198  
 Db 699 GTACGTCAAAACAAACGCAACGAGCGGACGGTCTCCGTGAGTGAGCATTTTAAAAAATGG 758  
 QY 1199 GAAAACTTAGGATGAATATGCGGAAAAATGTAAGTCGCTTACTGTAGAGGCTAT 1258  
 Db 759 GAAAGCTTAGGATGCAATGCGAATGGAAAAATGTAAGCAAGCAATTAAGTAGAGGCTAC 818  
 QY 1259 CAAAGTAGCGGAGTCTAATGATATAGCAATACATAAGAAATTAACGGTAAACCTCTC 1318  
 Db 819 CGAAGCAACGGAAGTGGGAATGTCATGACGAATCAGCTGATTCGATAAAAGCATATG 878  
 QY 1319 TCAACTATTAGTAAGCAGAGCAGCATACTTTGGATAAAAAACAATTAATAATCC 1372  
 Db 879 AAAAAAGCCAGCAAAAAATGGCTGCTTTTCTATGATAATTTTCAACTTCC 932

RESULT 13

US-08-275-526C-35  
 ; Sequence 35, Application US/08275526C  
 ; Patent No. 6180382  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DE BUYL, ERIC  
 ; APPLICANT: LAHAYE, ANDR E  
 ; APPLICANT: LEDOUX, PIERRE  
 ; APPLICANT: AMORY, ANTOINE  
 ; APPLICANT: DETROZ, REN  
 ; APPLICANT: ANDRE, CHRISTOPHE  
 ; APPLICANT: VETTER, ROMAN  
 ; TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,  
 ; TITLE OF INVENTION: EXPRESSION VECTORS FOR SUCH XYLANASE AND  
 ; TITLE OF INVENTION: OTHER PROTEINS, HOST ORGANISMS THEREFOR AND  
 ; TITLE OF INVENTION: USE THEREOF  
 ; NUMBER OF SEQUENCES: 35  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.  
 ; STREET: 2000 K St., N.W., Suite 200  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20006

COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/275,526C  
 ; FILING DATE: 15-JUL-1994  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Gadiano, Wilhem F.  
 ; REGISTRATION NUMBER: 37,136  
 ; REFERENCE/DOCKET NUMBER: 4121-49  
 ; TELEPHONE: (202) 429-0625  
 ; TELEFAX: (202) 293-0625  
 ; TELEX: 650 383 5605  
 ; INFORMATION FOR SEQ ID NO: 35:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1022 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: genomic DNA  
 ; US-08-275-526C-35

Query Match 18.7%; Score 282.8; DB 3; Length 1022;  
 Best Local Similarity 62.1%; Pred. No. 3.4e-66;  
 Matches 481; Conservative 0; Mismatches 287; Indels 6; Gaps 2;

QY 599 AATTAAGAGGAGGAGTGCCTAATGAGACAAAGAAATTTGACGTTGATTTTACCCCTTTTAA 658  
 Db 165 AAAAAAGAGAGGAGTGAATAATGAAATTTGAAAGATTGAGCTGTTGTTGTGATGTGT 224  
 QY 659 GTTTGTTTGGTCACTAACCTTACCTGCGAATAATTTAGGACACAATCGTCAACCCACAT 718  
 Db 225 ATTGATTTGCTGACACTGACGCTGTCGCGCTCATGCGGAAACGATTTATGATATAT 284  
 QY 719 TCCATTGCAACACGATGGCTATGATTTGAAATTTTGGAAAGATAGCGGTGGCTCTGG 778  
 Db 285 AGGATAGGAGACACACGCGGATACGATTTGAAATTTAGAGGATATC---GGAATACC 341  
 QY 779 ACAATGATTTCAATCATGCGGTAAGTTCAGTGCCTTCAATGGAACAAATGTTTAAACATA 838  
 Db 342 TCGATGACACTCAATAACGCGGGGCAATTTAGTGAAGCTGGAACAAATATTGGAATGCC 401  
 QY 839 TTATTCGTAAGGTAAAAATTTCAATGAAACACAAACACACCAACCAAGTTCGTAACATG 898  
 Db 402 TTATTTGAAAGGAAAGAGTTTGAATTCCTCACTAAATCTCATCATCACTTGGCAATC 461  
 QY 899 TCCATAAACTAGGAGCCAACTTCCAAACCAATGTAATGCGTATTTATGCGTCTATGTT 958  
 Db 462 TCCATCAACTACACGAGCCCTTTAAACCGGGCGGAATTCCTATTTATGTTCTATGGC 521  
 QY 959 TGGACTGTTGACCCCTTTTGTGAAATATTATATTGTCAGAGTTGGGCAACTGGCGTCCA 1018  
 Db 522 TGGACACAATCTCCATTAGCTGAATACATTGTTGAGTCATGCGGCACATATCGTCCA 581  
 QY 1019 CCAGGAGCAACGCTTAAGGGACCACTACTGTTGATGAGGAGCAATATGATATCTACGAG 1078  
 Db 582 ACAGG---AACGTATAAAGGATCAATTTATGCGGATGGAGGACATATGACATATATGAA 638  
 QY 1079 ACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTGCCACATTTAAACAATATTGGAGT 1138  
 Db 639 ACGCTCCGTGCAATCAGCCTTCTATCATTTGAGAGCGTACCTTCAACCAATATTGGAGT 698  
 QY 1139 GTTCGAAGATCGAAACGACGAGTGGCAGATTTCGTGAGCAACCACTTTAGAGCGTGG 1198  
 Db 699 GTACGTCACAAACAAACGACAGCGGAGCTCCGTCAGTGAGCATTTTAAAAAATGG 758  
 QY 1199 GAAAACTTAGGATGAATATGGGAAAAATGTAAGTCGCGCTTACTGTAGAGGCTAT 1258  
 Db 759 GAAAGCTTAGGATGCCAATGGGAAAAATGTAAGCAACAGCAATTAACCTGAGAGCTAC 818  
 QY 1259 CAAAGTAGCGGAGTGAATGATATAGCAATACACTAAGAAATTAACGGAACCTCTC 1318  
 Db 819 CAAAGCAACGGAAGTGGGAATGTCATGAGCAATCAGCTGATTCGATAAAAGCATATG 878  
 QY 1319 TCAACTATTAGTAATGACGAGAGCATACTTTGGATAAAAAACAATTAATAATCC 1372  
 Db 879 AAAAAAGCCAGCAAAAAATGGCTGCTTTTCTATGATAATTTTCAACTTCC 932

RESULT 14

US-09-076-677-1  
 ; Sequence 1, Application US/09076677  
 ; Patent No. 6423523

GENERAL INFORMATION:

APPLICANT: DE BUYL, ERIC  
 LAHAYE, ANDREE  
 LEDOUX, PIERRE  
 AMORY, ANTOINE  
 DETROZ, RENE  
 ANDRE, CHRISTOPHE  
 VETTER, ROMAN

TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,  
 EXPRESSION VECTORS FOR SUCH XYLANASE AND  
 OTHER PROTEINS, HOST ORGANISMS THEREFOR AND  
 USE THEREOF

NUMBER OF SEQUENCES: 35  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.  
 STREET: 2000 K St., N.W., Suite 200







US-09-076-677-35

Query Match		18.7%	Score 282.8;	DB 4;	Length 1022;
Best Local Similarity		62.1%	Pred. No. 3.4e-66;		
Matches 481;		Conservative 0;	Mismatches 287;	Indels 6;	Gaps 2;
Qy	599	AATTAAGAGGAGGAATGCCTAATGACACAAAGAAATTCAGCGTTGATTTTAGCCCTTTTAA	658		
Db	165	AAAAAGGAGAGGATGGAATGAATTTGAAGAATGAGGCTGTTGTTGTGATGCT	224		
Qy	659	GTGTTGTTTGCACCTTACCTGCGAGAAATTAATTCAGGCACAAATCGTCACCGACAAT	718		
Db	225	ATTGGATTTGTGCTGACACTGACGGCTGTCGGCTCATGCGGAACGATTTATGATAAT	284		
Qy	719	TCCATTGGCACAACGATGGCTATGATTAATTTTGGAAAGATAGCGGTGCTCTGGG	778		
Db	285	AGGATAGGGACACACAGCGGATACGATTTTGAATTTATGGAAGGATTAC---GGAAATACC	341		
Qy	779	ACAATGATTCTCAATCATGGCGGTAGCTTCAGTGCCTCAATGGAACAATGTTAAACAACATA	838		
Db	342	TCGATGACACTCAATTAACGGCGGGGCAATTTAGTGAAGCTGGAAACAATTTGGAATGCC	401		
Qy	839	TTATTCGCTAAGGTAAATAATCAATGAACAACAACACCAACAAGTTGGTAACATG	898		
Db	402	TTATTTGGAAGGAAAGAGTTTGAATCCACTAAACTCATCATCAACTTGGCAACATC	461		
Qy	899	TCCATAAATACGGAGCCAACTTCCAAACCAATGGTAATCGGTATTTATGCGTCTATGGT	958		
Db	462	TCCATCAACTACAAACGAGCCTTTAAACCGCGGGGGAATTCCTATTTATGTGTCTATGGC	521		
Qy	959	TGGACTGTTGACCTCTGTGCAATATTATTTGTGACAGATTGGGGCAACTGGCGTCCA	1018		
Db	522	TGGACAATCTCCATTAGCTGAATCTACTATTTGTTGAGTCATGCGGCACATATCTGCCA	581		
Qy	1019	CCAGGAGCAACGGCTTAAGGGGACCATCACTGTTGATGGAGGAACATATGATATCTACGAG	1078		
Db	582	ACAGG---AACGTATAAGGATCAATTTATGCCGATGGAGGCACATATGACATATATGAA	638		
Qy	1079	ACTCTTAGAGTCAATCAACCTCCATTAAAGGGATTCACACATTTAAACAATTTGGAGT	1138		
Db	639	ACGCTCGTGTCAATCAGCCTTCTATCATTTGGAGACGCTACCTTCAAAACAATATGGAGT	698		
Qy	1139	GTTCGAGATCGAAACGACGAGTGGCACGATTTCTGTGACCAACCACTTTAGAGCGTGG	1198		
Db	699	GTACGTCAAAACAACCCACAAGCGGAACGGTCTCCGTCACTGAGCAATTTAAAAATGG	758		
Qy	1199	GAAAACTTAGGGATGAATATGGGAAAATGTATGAAGTCGCGCTTACTGTAGAAGGCTAT	1258		
Db	759	GAAAGCTTAGGCATGCCAATGGGAAAATGTATGAACAGCATTTAACTGTAGAAGGCTAC	818		
Qy	1259	CAAAGTAGCGGAGTGTATATGATATATAGCAATACACTAGAATTAACGGTAACCTCTC	1318		
Db	819	CGAAGCAACGGAAGTGGCAATGTCAATGACGAATCAGCTGATGATTCGATAAAAGCATATG	878		
Qy	1319	TCACTATTAGTATAGGAGGACATAACTTTGGATAAAACAATTAATAATCC	1372		
Db	879	AAAAAGCCGCAAAAATGGCTGGCTTTTCTATGATAATTTTTCACACTTC	932		

Search completed: November 11, 2004, 01:34:23  
Job time : 176.81 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 21:26:43 ; Search time 967.351 seconds  
(without alignments)  
8421.106 Million cell updates/sec

Title: US-09-909-207-10  
Perfect score: 1513  
Sequence: 1 ARAATGAATTGGTATATCT.....TGAACACCTCGTCACTAG 1513

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3611042 seqs, 2692057975 residues  
Total number of hits satisfying chosen parameters: 7222084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1513	100.0	1513	9	US-09-909-207-10
2	1513	100.0	1513	9	US-09-909-207-11
3	744	49.2	744	9	US-09-909-207-4
4	744	49.2	744	9	US-09-909-207-5
5	663	43.8	663	9	US-09-909-207-1
6	663	43.8	663	9	US-09-909-207-2
7	619	40.9	619	9	US-09-909-207-12
8	150	9.9	150	9	US-09-909-207-13
9	119.4	7.9	1375	15	US-09-770-621-1
10	119.4	7.9	1375	15	US-10-286-993-1
11	115.2	7.6	596	15	US-10-307-441-39
12	111.4	7.4	942	14	US-10-213-990-71
					Sequence 10, Appl
					Sequence 11, Appl
					Sequence 4, Appl
					Sequence 5, Appl
					Sequence 1, Appl
					Sequence 2, Appl
					Sequence 12, Appl
					Sequence 13, Appl
					Sequence 1, Appl
					Sequence 39, Appl
					Sequence 71, Appl

13	110	7.3	1002	14	US-10-213-990-70	Sequence 70, Appl
14	108.8	7.2	705	14	US-10-213-990-68	Sequence 68, Appl
15	106	7.0	2898	15	US-10-299-393-1	Sequence 1, Appl
16	104.8	6.9	712	14	US-10-213-990-64	Sequence 64, Appl
17	103.6	6.8	666	14	US-10-213-990-65	Sequence 65, Appl
18	96.8	6.4	739	14	US-10-213-990-67	Sequence 67, Appl
19	86.8	5.7	678	10	US-09-803-454-3	Sequence 3, Appl
20	84.2	5.6	1023	16	US-10-244-596-6	Sequence 6, Appl
21	84	5.6	1027	16	US-10-244-596-10	Sequence 10, Appl
22	83.2	5.5	645	15	US-10-237-386-11	Sequence 11, Appl
23	83.2	5.5	657	15	US-10-237-386-11	Sequence 11, Appl
24	82.6	5.5	1011	16	US-10-244-596-12	Sequence 12, Appl
25	82.6	5.5	1011	16	US-10-244-596-13	Sequence 13, Appl
26	82.6	5.5	1023	16	US-10-244-596-5	Sequence 5, Appl
27	82.6	5.5	1023	16	US-10-244-596-7	Sequence 7, Appl
28	82.6	5.5	1023	16	US-10-244-596-8	Sequence 8, Appl
29	82.6	5.5	1023	16	US-10-244-596-9	Sequence 9, Appl
30	82.6	5.5	1023	16	US-10-244-596-9	Sequence 9, Appl
31	81	5.4	81	9	US-09-909-207-7	Sequence 7, Appl
32	81	5.4	81	9	US-09-909-207-8	Sequence 8, Appl
33	81	5.4	1023	16	US-10-244-596-3	Sequence 3, Appl
34	81	5.4	1023	16	US-10-244-596-10	Sequence 10, Appl
35	79.4	5.2	1023	16	US-10-244-596-4	Sequence 4, Appl
36	79	5.2	2225	10	US-03-790-070A-8	Sequence 8, Appl
37	77.8	5.1	1014	16	US-10-244-596-1	Sequence 1, Appl
38	77.2	5.1	749	18	US-10-425-115-82922	Sequence 82922, A
39	75.8	5.0	818	18	US-10-425-115-37173	Sequence 37173, A
40	73.8	4.9	850	18	US-10-425-115-177283	Sequence 177283, A
41	72.4	4.8	588	15	US-10-237-386-9	Sequence 9, Appl
42	72.4	4.8	983	9	US-09-467-368-1	Sequence 1, Appl
43	57.6	3.8	3673778	15	US-10-312-841-2	Sequence 2, Appl
44	55.8	3.7	3673778	15	US-10-312-841-1	Sequence 1, Appl
45	54.2	3.6	2054	15	US-10-419-969-5	Sequence 5, Appl

## ALIGNMENTS

RESULT 1  
US-09-909-207-10  
; Sequence 10, Application US/09909207  
; Patent No. US20020115181A1  
; GENERAL INFORMATION:  
; APPLICANT: ANDREE LAHAYE  
; ERIC DE BUYL  
; PIERRE LEDOUX  
; RENE DETROZ

TITLE OF INVENTION: Xylanase, microorganisms produced it,  
DNA molecule, processes for preparation of this xylanase  
and uses thereof

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE

STREET: 2000 K St., N.W., Suite 200

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

CURRENT APPLICATION DATA: US/09/909,207

APPLICATION NUMBER: 19-Jul-2001

FILING DATE: 19-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/470,953

FILING DATE: 06-JUNE-1995

ATTORNEY/AGENT INFORMATION:

NAME: Wilhem F. Gadiano, Esq.

REGISTRATION NUMBER: 37,136

REFERENCE/DOCKET NUMBER: 4121-40  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1513 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: Bacillus  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-09-909-207-10

Query Match 100.0%; Score 1513; DB 9; Length 1513;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATTGAATTGTATCTAATGATAACGACAAATCGTCACTGTTTTTAAACTAATCTC 60  
DB 1 AAATTGAATTGTATCTAATGATAACGACAAATCGTCACTGTTTTTAAACTAATCTC 60

QY 61 AAACCAATCTCTTTTATTTAAGCTAACCACTTGCATCTTATCACAGAACATCTTT 120  
DB 61 AAACCAATCTCTTTTATTTAAGCTAACCACTTGCATCTTATCACAGAACATCTTT 120

QY 121 ATAGGAACCTTCCCATTTTCAAGACGATAAAAATCTTTTCCCTATTTTATCTATCG 180  
DB 121 ATAGGAACCTTCCCATTTTCAAGACGATAAAAATCTTTTCCCTATTTTATCTATCG 180

QY 181 CTTGATCGGTTTAAATTTGTAACCTTTATTTAGTTTACGTTGATGTTCCCTCATTCAC 240  
DB 181 CTTGATCGGTTTAAATTTGTAACCTTTATTTAGTTTACGTTGATGTTCCCTCATTCAC 240

QY 241 CATTAAATCAGTTAAGCTGAGTCACTCTTTTTCGTTCTCAAAATACCTGGAAC 300  
DB 241 CATTAAATCAGTTAAGCTGAGTCACTCTTTTTCGTTCTCAAAATACCTGGAAC 300

QY 301 ATTTATGTCATATTTCTCACGCGCTCCATATGAATATATATATCTTTTATACATA 360  
DB 301 ATTTATGTCATATTTCTCACGCGCTCCATATGAATATATATATCTTTTATACATA 360

QY 361 TTAAGTAAATAGTATATCTTGGGTTATCAAAATGTGAGATAATCTAATGTCAACA 420  
DB 361 TTAAGTAAATAGTATATCTTGGGTTATCAAAATGTGAGATAATCTAATGTCAACA 420

QY 421 AGCAGCTATCCAAAAAAGTATGTCATCTTAAAGAGTGCATCTATCTATGAAA 480  
DB 421 AGCAGCTATCCAAAAAAGTATGTCATCTTAAAGAGTGCATCTATCTATGAAA 480

QY 481 GATAATATCCAGTTTCAAAATTTGAAATAGTGTGTATGGAATAGTTTGAATGTCAA 540  
DB 481 GATAATATCCAGTTTCAAAATTTGAAATAGTGTGTATGGAATAGTTTGAATGTCAA 540

QY 541 CTGTGAAGGAGGTAGGTAGTACGTTAGTCTTCAATACCAAAATAGTTGTAAAAA 600  
DB 541 CTGTGAAGGAGGTAGGTAGTACGTTAGTCTTCAATACCAAAATAGTTGTAAAAA 600

QY 601 TTAAGAGGAGGATGCTAATCAGACAAAGAAATTTAGTTGATTTAGCTTTTGTAGT 660  
DB 601 TTAAGAGGAGGATGCTAATCAGACAAAGAAATTTAGTTGATTTAGCTTTTGTAGT 660

QY 661 TTGTTTTGCACTAACCTTACCTGAGAAATATTCAGGCACAAATCGTCAACGCAATTC 720  
DB 661 TTGTTTTGCACTAACCTTACCTGAGAAATATTCAGGCACAAATCGTCAACGCAATTC 720

QY 721 CATTGGCAACCGATGGCTATGATTAATTTTGGAAAGATAGCGGTGGCTCTGGGAC 780  
DB 721 CATTGGCAACCGATGGCTATGATTAATTTTGGAAAGATAGCGGTGGCTCTGGGAC 780

QY 781 AATGATTTCTCAATCATGCGGTACGTTTCCAGTCCCAATGGAACAATGTTAAACAATATT 840  
DB 781 AATGATTTCTCAATCATGCGGTACGTTTCCAGTCCCAATGGAACAATGTTAAACAATATT 840

QY 841 ATTCGGTAAAGGTAAAAAATTCATATGAACAACCAACACCAACCAAGTTGGTAACTGTC 900  
DB 841 ATTCGGTAAAGGTAAAAAATTCATATGAACAACCAACACCAACCAAGTTGGTAACTGTC 900

QY 901 CATAACTACGAGGCAACTTCCAAACCAATGGAATGCGTATTTATGCGTCTATGGTTG 960  
DB 901 CATAACTACGAGGCAACTTCCAAACCAATGGAATGCGTATTTATGCGTCTATGGTTG 960

QY 961 GACTGTTGACCTCTTGTGCAATATATTTATGTCAGAGTTGGGCAACTGCGTCCACC 1020  
DB 961 GACTGTTGACCTCTTGTGCAATATATTTATGTCAGAGTTGGGCAACTGCGTCCACC 1020

QY 1021 AGGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGAACATATGATATCTACGAGAC 1080  
DB 1021 AGGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGAACATATGATATCTACGAGAC 1080

QY 1081 TCTTAGAGTCAATCAACCCCTCATTAAGGGGATGGCACATTTAAACAATATTGGAGTGT 1140  
DB 1081 TCTTAGAGTCAATCAACCCCTCATTAAGGGGATGGCACATTTAAACAATATTGGAGTGT 1140

QY 1141 TCGAAGATCGAAACGACGAGTGGCAGGATTTCTGTCAGCAACCACTTTAGAGCGTGGGA 1200  
DB 1141 TCGAAGATCGAAACGACGAGTGGCAGGATTTCTGTCAGCAACCACTTTAGAGCGTGGGA 1200

QY 1201 AAACCTTAGGATGAATATGGGAAAAATGATGAAGTCCGCTTACTGTAGAGGCTATCA 1260  
DB 1201 AAACCTTAGGATGAATATGGGAAAAATGATGAAGTCCGCTTACTGTAGAGGCTATCA 1260

QY 1261 AAGTAGGGAAGTCTAATGTATATAGCAATACACTAAGAAATTAACCGTAAACCTCTCTC 1320  
DB 1261 AAGTAGGGAAGTCTAATGTATATAGCAATACACTAAGAAATTAACCGTAAACCTCTCTC 1320

QY 1321 AACTATTAGTAAATCAGCAGAGCATAAATTTGGATAAAAACAATTTAAAAATCCTTTCTCT 1380  
DB 1321 AACTATTAGTAAATCAGCAGAGCATAAATTTGGATAAAAACAATTTAAAAATCCTTTCTCT 1380

QY 1381 TCGGTTTCAAGTCTCATTTTCAAAATAACCTCCCGTTGGATCTTTTCCAAACGGGAGG 1440  
DB 1381 TCGGTTTCAAGTCTCATTTTCAAAATAACCTCCCGTTGGATCTTTTCCAAACGGGAGG 1440

QY 1441 TTTTATTTGGAAGGTTAAGTATAGTATCTCCGATTCATCCAGAGGATGCTTGAACA 1500  
DB 1441 TTTTATTTGGAAGGTTAAGTATAGTATCTCCGATTCATCCAGAGGATGCTTGAACA 1500

QY 1501 CCTCCGTCACCTAG 1513  
DB 1501 CCTCCGTCACCTAG 1513

RESULT 2  
US-09-909-207-11  
; Sequence 11, Application US/09909207  
; Patent No. US20020115181A1  
; GENERAL INFORMATION:  
; APPLICANT: ANDREE LAHAYE  
; ERIC DE BUYL  
; PIERRE LEDOUX  
; RENE DETROZ  
; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
; DNA molecule, processes for preparation of this xylanase  
; and uses thereof  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
; STREET: 2000 K St., N.W., Suite 200  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/909,207
FILING DATE: 19-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/470,953
FILING DATE: 06-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Wilhem F. Gadiano, Esq.
REGISTRATION NUMBER: 37,136
REFERENCE/DOCKET NUMBER: 4121-40
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-429-0625
TELEFAX: (202) 293-1850
TELEX: 650 383-5605
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1513 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
STRAIN: Bacillus
FEATURE:
NAME/KEY: CDS
LOCATION: 620..1363
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 701..1363
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 620..700
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-909-207-11

```

```

Query Match      100.0%; Score 1513; DB 9; Length 1513;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATTGAATTGTATATCTAATGATTAACGACAAATCGTCACCTGTTTAAACATATCTC 60
DB 1 AAATTGAATTGTATATCTAATGATTAACGACAAATCGTCACCTGTTTAAACATATCTC 60

QY 61 AAACCAATCTCTTTTAAATTAACGCTAACCACTTGAATCTTATCAAGAAACATCTTT 120
DB 61 AAACCAATCTCTTTTAAATTAACGCTAACCACTTGAATCTTATCAAGAAACATCTTT 120

QY 121 ATAGGAATCTTCCCAATTTGCAAGACGATAAAAAATCTTTTCCCTATTTTATCTATCG 180
DB 121 ATAGGAATCTTCCCAATTTGCAAGACGATAAAAAATCTTTTCCCTATTTTATCTATCG 180

QY 181 CTTTGATCGGTTAAATTTGTAACCTTTTATTTAGTTTACGTGATGTTCCCTCATTCATAC 240
DB 181 CTTTGATCGGTTAAATTTGTAACCTTTTATTTAGTTTACGTGATGTTCCCTCATTCATAC 240

QY 241 CATTAATCACAGTTAAACGCTAGAGTCATCTTTTCCGTTCTCAAAATACCTGAAAGAC 300
DB 241 CATTAATCACAGTTAAACGCTAGAGTCATCTTTTCCGTTCTCAAAATACCTGAAAGAC 300

QY 301 ATTTATGTCATATTTCTCAGCGCGCTCCATAATGGAATATATATCTCTTTTATACATA 360
DB 301 ATTTATGTCATATTTCTCAGCGCGCTCCATAATGGAATATATATCTCTTTTATACATA 360

QY 361 TTAAGTAAATTAGTATATCTTGGGTTATCAAAATGTGAGATATCTTAATGTGATCAAA 420
DB 361 TTAAGTAAATTAGTATATCTTGGGTTATCAAAATGTGAGATATCTTAATGTGATCAAA 420

```

```

QY 421 AGCAGCTATCCAAAAAACACATGATGTTGACCTCTTTAAAGAAAGTGTCACATCTATGA 480
DB 421 AGCAGCTATCCAAAAAACACATGATGTTGACCTCTTTAAAGAAAGTGTCACATCTATGA 480

QY 481 GATAATATCCAGTTTCAAAATTTGAATAGTGTATGGAATAGCTTTGAATGTCACTG 540
DB 481 GATAATATCCAGTTTCAAAATTTGAATAGTGTATGGAATAGCTTTGAATGTCACTG 540

QY 541 CTGTGAAAGGAGGAGGTAGTAGTACCGTAGACTTCAATTACCAAAAATTTAGTTGTAA 600
DB 541 CTGTGAAAGGAGGAGGTAGTAGTACCGTAGACTTCAATTACCAAAAATTTAGTTGTAA 600

QY 601 TTAAGGAGGAGGAGGTAGTACCGTAGACTTCAATTACCAAAAATTTAGTTGTAA 660
DB 601 TTAAGGAGGAGGAGGTAGTACCGTAGACTTCAATTACCAAAAATTTAGTTGTAA 660

QY 661 TTGTTTTCGACTTAACCTTACCTGCGAGAAATTAATTCAGGCACAAATCGTCACGACA 720
DB 661 TTGTTTTCGACTTAACCTTACCTGCGAGAAATTAATTCAGGCACAAATCGTCACGACA 720

QY 721 CATTCGCAACACGATGCTATGATTATCAATTTTGGAAAGATAGCGGTGGCTCTGGGAC 780
DB 721 CATTCGCAACACGATGCTATGATTATCAATTTTGGAAAGATAGCGGTGGCTCTGGGAC 780

QY 781 AATGATTTCTCAATCATGCGGTACCTTCAGTCCCAATGGAACAATGTTTAAACAATAT 840
DB 781 AATGATTTCTCAATCATGCGGTACCTTCAGTCCCAATGGAACAATGTTTAAACAATAT 840

QY 841 ATTCGTTAAAGGTAAGAAATTTCAATGAAACACAAACACCAACCAAGTTGGTAAATGTC 900
DB 841 ATTCGTTAAAGGTAAGAAATTTCAATGAAACACAAACACCAACCAAGTTGGTAAATGTC 900

QY 901 CATAACTACGAGGACCACTTCCAAACCAATGCTAATGCTAATTTATGCGTCTATGGTTG 960
DB 901 CATAACTACGAGGACCACTTCCAAACCAATGCTAATGCTAATTTATGCGTCTATGGTTG 960

QY 961 GACTGTGACCTCTTGTGCAATATATTTGTCAGAGTGTGGGCAATCGCGTCCACC 1020
DB 961 GACTGTGACCTCTTGTGCAATATATTTGTCAGAGTGTGGGCAATCGCGTCCACC 1020

QY 1021 AGGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGAAACATATGATATCTACGAGAC 1080
DB 1021 AGGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGAAACATATGATATCTACGAGAC 1080

QY 1081 TCTTAGAGTCAATCAACCTCCATTAAGGGGATGCGCAATTTAAACAATATTTGGAGTGT 1140
DB 1081 TCTTAGAGTCAATCAACCTCCATTAAGGGGATGCGCAATTTAAACAATATTTGGAGTGT 1140

QY 1141 TCGAAGATCGAAACGACGAGTGGCAGATTTCTGTACGACCAACCACTTTAGAGCGTGGGA 1200
DB 1141 TCGAAGATCGAAACGACGAGTGGCAGATTTCTGTACGACCAACCACTTTAGAGCGTGGGA 1200

QY 1201 AAACCTTAGGAGTGAATATGGGGAATGATGAGTCCGCTTACTGTAGAGGCTATCA 1260
DB 1201 AAACCTTAGGAGTGAATATGGGGAATGATGAGTCCGCTTACTGTAGAGGCTATCA 1260

QY 1261 AAGTACGGAAGTGTAAATGATATAGCAATACACCTAAGAAATTAACGGTAAACCTCTCTC 1320
DB 1261 AAGTACGGAAGTGTAAATGATATAGCAATACACCTAAGAAATTAACGGTAAACCTCTCTC 1320

QY 1321 AACTATTAGTATGACGAGACGATTAACCTTTGGATAAAAACAATTAATAAATCTTTATCTCT 1380
DB 1321 AACTATTAGTATGACGAGACGATTAACCTTTGGATAAAAACAATTAATAAATCTTTATCTCT 1380

QY 1381 TTGGTTTCAAGTTCTCATTTATTTCAATAACCTCCGGTGGATCTTTTCCACGCGGAGG 1440
DB 1381 TTGGTTTCAAGTTCTCATTTATTTCAATAACCTCCGGTGGATCTTTTCCACGCGGAGG 1440

QY 1441 TTTTATTGGAAGGTTAAGTATAGTATCTCCGATTTCCATCCAGAGGAATGCTTTGAACA 1500
DB 1441 TTTTATTGGAAGGTTAAGTATAGTATCTCCGATTTCCATCCAGAGGAATGCTTTGAACA 1500

QY 1501 CCTCGCTCACTAG 1513

```

Db 1501 CCTCGTCACTAG 1513

RESULT 3  
US-09-909-207-4  
; Sequence 4, Application US/09909207  
; Patent No. US20020115181A1  
; GENERAL INFORMATION:  
; APPLICANT: ANDREE LAHAYE  
; ERIC DE BOYL  
; PIERRE LEDOUX  
; RENE DETROZ  
; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
; DNA molecule, processes for preparation of this xylanase  
; and uses thereof  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
; STREET: 2000 K St., N.W., Suite 200  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/09/909,207  
; FILING DATE: 19-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/470,953  
; FILING DATE: 06-JUNE-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wilhem F. Gadiano, Esq.  
; REGISTRATION NUMBER: 37,136  
; REFERENCE/DOCKET NUMBER: 4121-40  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-429-0625  
; TELEFAX: (202) 293-1850  
; TELEX: 650 383-5605  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 744 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; STRAIN: Bacillus  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-909-207-4  
Query Match 49.2%; Score 744; DB 9; Length 744;  
Best Local Similarity 100.0%; Pred. No. 1.3e-166;  
Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 620 ATGACAAAAGAAATGACGTTGATTTAGCCTTTTGTGTTGTTGCTACTAACCTTA 679  
Db 1 ATGACAAAAGAAATGACGTTGATTTAGCCTTTTGTGTTGTTGCTACTAACCTTA 60  
QY 680 CCTGCAGAAATATTCAGGCAAAATCGTCACCGCAATTCATTTGGCAACACGATGC 739  
Db 61 CCTGCAGAAATATTCAGGCAAAATCGTCACCGCAATTCATTTGGCAACACGATGC 120  
QY 740 TATGATTATGAATTTTGGAAAGATAGCGTGGCTCTGGGACAAATGATTCTCAATCATGGC 799  
Db 121 TATGATTATGAATTTTGGAAAGATAGCGTGGCTCTGGGACAAATGATTCTCAATCATGGC 180  
QY 800 GGTACGTTCACTGCGCCCAATGGAAACAATGTTTAAACAACATATTATTCCTGTAAGGTAAGAAA 859

Db 181 GGTACGTTCACTGCGCCCAATGGAAACAATGTTTAAACAACATATTATTCCTGTAAGGTAAGAAA 240  
QY 860 TTCAATGAACACAAACACACACCAACAAAGTTGGTGAACATGTCCTCAATAAATCTACGAGCCCAAC 919  
Db 241 TTCAATGAACACAAACACACACCAACAAAGTTGGTGAACATGTCCTCAATAAATCTACGAGCCCAAC 300  
QY 920 TTCCAACCAATGTAATGCTTATTTATGCTCTCTGTTGGACTGTTGACCTCTTGTGTC 979  
Db 301 TTCCAACCAATGTAATGCTTATTTATGCTCTCTGTTGGACTGTTGACCTCTTGTGTC 360  
QY 980 GAATATTATATTGTCGACAGTTGGGCAACTGGCGTCCACACAGAGCAACCGCTTAAGGGG 1039  
Db 361 GAATATTATATTGTCGACAGTTGGGCAACTGGCGTCCACACAGAGCAACCGCTTAAGGGG 420  
QY 1040 ACCATCACTGTTGATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 1099  
Db 421 ACCATCACTGTTGATGAGGAAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCC 480  
QY 1100 TCCATTAAAGGGATTGCCACATTTTAAACAATATTGGAGTGTTCGAAGATCGAAACGCAACG 1159  
Db 481 TCCATTAAAGGGATTGCCACATTTTAAACAATATTGGAGTGTTCGAAGATCGAAACGCAACG 540  
QY 1160 AGTGGCAGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAACTTTAGGATGAATATG 1219  
Db 541 AGTGGCAGATTTCTGTGACCAACCACTTTAGAGCGTGGGAAACTTTAGGATGAATATG 600  
QY 1220 GGGAAATGATGAAGTCGGCTTACTGTAGAGGCTTCAAGTAGGGAGTGCCTAAT 1279  
Db 601 GGGAAATGATGAAGTCGGCTTACTGTAGAGGCTTCAAGTAGGGAGTGCCTAAT 660  
QY 1280 GTATATAGCAATACACTAAGAAATTAACGTAACCTCTCTCAACTATTAGTAATGACGAG 1339  
Db 661 GTATATAGCAATACACTAAGAAATTAACGTAACCTCTCTCAACTATTAGTAATGACGAG 720  
QY 1340 AGCAATACTTTGGATAAAACAAT 1363  
Db 721 AGCAATACTTTGGATAAAACAAT 744

RESULT 4  
US-09-909-207-5  
; Sequence 5, Application US/09909207  
; Patent No. US20020115181A1  
; GENERAL INFORMATION:  
; APPLICANT: ANDREE LAHAYE  
; ERIC DE BOYL  
; PIERRE LEDOUX  
; RENE DETROZ  
; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
; DNA molecule, processes for preparation of this xylanase  
; and uses thereof  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
; STREET: 2000 K St., N.W., Suite 200  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/09/909,207  
; FILING DATE: 19-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/470,953  
; FILING DATE: 06-JUNE-1995  
; ATTORNEY/AGENT INFORMATION:

NAME: Wilhem F. Gadiano, Esq.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-40  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-0625  
TELEFAX: (202) 293-1850  
TELEX: 650 383-5605  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 744 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
STRAIN: Bacillus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..744  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 82..744  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 1..81  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-909-207-5  
Query Match 49.2%; Score 744; DB 9; Length 744;  
Best Local Similarity 100.0%; Pred. No. 1.3e-166;  
Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 620 ATGAGACAAAGAAATGACGTTGATTTAGCTTTTATGCTTTTATGTTGTTGTTGCTTAACCTTA 679  
DB 1 ATGAGACAAAGAAATGACGTTGATTTAGCTTTTATGCTTTTATGTTGTTGTTGCTTAACCTTA 60  
QY 680 CTGCGAGAATAATTCAGGCACAAATCGTCACGCAATTCATTCGACCAACCAACGATGC 739  
DB 61 CTGCGAGAATAATTCAGGCACAAATCGTCACGCAATTCATTCGACCAACCAACGATGC 120  
QY 740 TATGATTATCAATTTTGGAAAGATAGCGTGGCTCTGGGACCAATGATTTCTCAATCATGCG 799  
DB 121 TATGATTATCAATTTTGGAAAGATAGCGTGGCTCTGGGACCAATGATTTCTCAATCATGCG 180  
QY 800 GGTACGTTACGTGCCCAATGGAAATGTTTAAACAATATATTCGTTAAAGTAAAAA 859  
DB 181 GGTACGTTACGTGCCCAATGGAAATGTTTAAACAATATATTCGTTAAAGTAAAAA 240  
QY 860 TTCATGAACACAAACACCAACCAAGTTGGTAAATGTCATGTCCTCCATAACTACGGAGCCAAC 919  
DB 241 TTCATGAACACAAACACCAACCAAGTTGGTAAATGTCATGTCCTCCATAACTACGGAGCCAAC 300  
QY 920 TTCACCAATATGTAATGCTATTTATGCTCTATGTTGGACTGTGACCTCTTTGTC 979  
DB 301 TTCACCAATATGTAATGCTATTTATGCTCTATGTTGGACTGTGACCTCTTTGTC 360  
QY 980 GAATATTTATTTGTCAGTTGGGCAACTGGCGTCCACGAGCAACGCTTAAGGGG 1039  
DB 361 GAATATTTATTTGTCAGTTGGGCAACTGGCGTCCACGAGCAACGCTTAAGGGG 420  
QY 1040 ACATCATCTGTTGATGAGGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACC 1099  
DB 421 ACCATCATCTGTTGATGAGGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACC 480  
QY 1100 TCCATTAAGGGGATTCGCACATTTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGCACG 1159  
DB 481 TCCATTAAGGGGATTCGCACATTTTAAACAATATTTGGAGTGTTCGAAGATCGAAACGCACG 540  
QY 1160 AGTGGCAGATTTCTGTCAGCAACCACTTTTAGAGCGTGGGAAAATTAGGGATGAATATG 1219  
DB 541 AGTGGCAGATTTCTGTCAGCAACCACTTTTAGAGCGTGGGAAAATTAGGGATGAATATG 600  
QY 1220 GGGAAATATGTAAGTGGCGCTTACTGTAGAGGCTATCAAGTAGCGGAAAGTCTAAT 1279

DB 601 GGGAAATATGTAAGTGGCGCTTACTGTAGAGGCTATCAAGTAGCGGAAGTGTCTAAT 660  
QY 1280 GTATATAGCAATACACTAAGAAATTAACGGTAACCTCTCTCAACTATTAGTAATGACGAG 1339  
DB 661 GTATATAGCAATACACTAAGAAATTAACGGTAACCTCTCTCAACTATTAGTAATGACGAG 720  
QY 1340 AGCATAACTTTGGATAAAAAACAAT 1363  
DB 721 AGCATAACTTTGGATAAAAAACAAT 744  
RESULT 5  
US-09-909-207-1  
; Sequence 1, Application US/09909207  
; Patent No. US20020115181A1  
; GENERAL INFORMATION:  
; APPLICANT: ANDREE LAHAYE  
; ERIC DE BUYL  
; PIERRE LEDOUX  
; RENE DETROZ  
; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
; DNA molecule, processes for preparation of this xylanase  
; and uses thereof  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE  
; STREET: 2000 K St., N.W., Suite 200  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/909,207  
; FILING DATE: 19-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/470,953  
; FILING DATE: 06-JUNE-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wilhem F. Gadiano, Esq.  
; REGISTRATION NUMBER: 37,136  
; REFERENCE/DOCKET NUMBER: 4121-40  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-429-0625  
; TELEFAX: (202) 293-1850  
; TELEX: 650 383-5605  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 663 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; STRAIN: Bacillus  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-909-207-1  
Query Match 43.8%; Score 663; DB 9; Length 663;  
Best Local Similarity 100.0%; Pred. No. 2.3e-147;  
Matches: 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 701 CAAATCGTCCACCAATTCCTTGGCAACCAACGATGGCTATGATTATGATTTTGGAAA 760  
DB 1 CAAATCGTCCACCAATTCCTTGGCAACCAACGATGGCTATGATTATGATTTTGGAAA 60  
QY 761 GATAGCGGTGGCTCTGGGACCAATGATTTCTCAATCATGGCGGTACGTTAGTCCCAATGG 820





QY 1301 ATTAACGGTAACCTCTCTCAACTATTAGTAATGACGAGACATACCTTTGGATATAAAC 1360  
 DB 601 ATTAACGGTAACCTCTCTCAACTATTAGTAATGACGAGACATACCTTTGGATATAAAC 660  
 QY 1361 AAT 1363  
 DB 661 AAT 663

RESULT 7

US-09-909-207-12  
 ; Sequence 12, Application US/09909207  
 ; Patent No. US20020115181A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ANDREE LAHAYE  
 ; ERIC DE BUYL  
 ; PIERRE LEDOUX  
 ; RENE DETROZ  
 ; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
 ; DNA molecule, processes for preparation of this xylanase  
 ; and uses thereof  
 ;  
 ; NUMBER OF SEQUENCES: 29  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: WILLIAN BRINKS HOPFER GILSON & LIONE  
 ; STREET: 2000 K St., N.W., Suite 200  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20006  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/909,207  
 ; FILING DATE: 19-Jul-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/470,953  
 ; FILING DATE: 06-JUNE-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Wilhem F. Gadiano, Esq.  
 ; REGISTRATION NUMBER: 37,136  
 ; REFERENCE/DOCKET NUMBER: 4121-40  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-429-0625  
 ; TELEFAX: (202) 231-1850  
 ; TELEX: 650 383-5605  
 ; INFORMATION FOR SEQ ID NO: 12:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 619 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Query Match 40.9%; Score 619; DB 9; Length 619;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-137;  
 Matches 619; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AAATTGAATTCGTATATCTATGATACGACCAATCGTCACTGTTTTTAACTAATCTC 60  
 DB 1 AAATTGAATTCGTATATCTATGATACGACCAATCGTCACTGTTTTTAACTAATCTC 60  
 QY 61 AAACCAATACCTCTTTATTTAACTGACCACTGCAATCTTATACAGAACATCTTTT 120  
 DB 61 AAACCAATACCTCTTTATTTAACTGACCACTGCAATCTTATACAGAACATCTTTT 120  
 QY 121 ATAGAACTTCCCAATTTGGACGACGATAAAATCTTTTCCCTATTATCTTATCG 180

DB 121 ATAGAACTTCCCAATTTGGACGACGATAAAATCTTTTCCCTATTATCTTATCTG 180  
 QY 181 CCTTGATCGGTTTAAATTTGTAACCTTTTATTTAGTTTACGTCATGTTCCCTCATTCATAC 240  
 DB 181 CCTTGATCGGTTTAAATTTGTAACCTTTTATTTAGTTTACGTCATGTTCCCTCATTCATAC 240  
 QY 241 CATTAAATCACAGTTTAAACGCTAGAGTCATCTTTTTCGGTTCTCAAAAATACCTGAAGAAC 300  
 DB 241 CATTAAATCACAGTTTAAACGCTAGAGTCATCTTTTTCGGTTCTCAAAAATACCTGAAGAAC 300  
 QY 301 ATTTATGTCATATTTTCTCAGCCGCTCCATAATGGAATATATATCTTTTATACATA 360  
 DB 301 ATTTATGTCATATTTTCTCAGCCGCTCCATAATGGAATATATATCTTTTATACATA 360  
 QY 361 TTAAGTAAATTAGTATATATCTTGCCTTATCAAAAATGTCAGATAATCTTAATTCATCAACA 420  
 DB 361 TTAAGTAAATTAGTATATATCTTGCCTTATCAAAAATGTCAGATAATCTTAATTCATCAACA 420  
 QY 421 AGCAGCTATCCAAAAACACATGATGTTGACCTCTTTAAAGAGAGTGTCACTATCTATGAAAA 480  
 DB 421 AGCAGCTATCCAAAAACACATGATGTTGACCTCTTTAAAGAGAGTGTCACTATCTATGAAAA 480  
 QY 481 GATAATTTATCCAGTTTCAAAAATTTGAAATAGTGTGATGGAATAGTTTGAATGTCACTG 540  
 DB 481 GATAATTTATCCAGTTTCAAAAATTTGAAATAGTGTGATGGAATAGTTTGAATGTCACTG 540  
 QY 541 CTGTGAAAGGAGGTTAGTAGTACCGTAGCTTCACTTACCAAAAATTTAGTTGTGAAAAAAA 600  
 DB 541 CTGTGAAAGGAGGTTAGTAGTACCGTAGCTTCACTTACCAAAAATTTAGTTGTGAAAAAAA 600  
 QY 601 TTAAGAGGAGGAGTGCCTA 619  
 DB 601 TTAAGAGGAGGAGTGCCTA 619

RESULT 8

US-09-909-207-13  
 ; Sequence 13, Application US/09909207  
 ; Patent No. US20020115181A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ANDREE LAHAYE  
 ; ERIC DE BUYL  
 ; PIERRE LEDOUX  
 ; RENE DETROZ  
 ; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
 ; DNA molecule, processes for preparation of this xylanase  
 ; and uses thereof  
 ;  
 ; NUMBER OF SEQUENCES: 29  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: WILLIAN BRINKS HOPFER GILSON & LIONE  
 ; STREET: 2000 K St., N.W., Suite 200  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20006  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/909,207  
 ; FILING DATE: 19-Jul-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/470,953  
 ; FILING DATE: 06-JUNE-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Wilhem F. Gadiano, Esq.  
 ; REGISTRATION NUMBER: 37,136  
 ; REFERENCE/DOCKET NUMBER: 4121-40  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-429-0625

```

; TELEFAX: (202) 293-1850
; TELEX: 650 383-5605
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-909-207-13

Query Match
Best Local Similarity 9.9%; Score 150; DB 9; Length 150;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1364 TAAATATCTTATCTCTTTCGGTTCAGTTCATTTTCAATTAACCTCCCGGTTGGA 1423
Db 1 TAAATATCTTATCTCTTTCGGTTCAGTTCATTTTCAATTAACCTCCCGGTTGGA 60

QY 1424 TCTTTTCCAAACGGGAGGTTTATTGGAAAGGTTAAGTATAGTATCTCGGATTCATCCA 1483
Db 61 TCTTTTCCAAACGGGAGGTTTATTGGAAAGGTTAAGTATAGTATCTCGGATTCATCCA 120

QY 1484 GAGGAATGCTTGAACACCTCGTCACTAG 1513
Db 121 GAGGAATGCTTGAACACCTCGTCACTAG 150

```

```

RESULT 9
US-09-770-621-1
; Sequence 1, Application US/09770621
; Patent No. US20010024815A1
; GENERAL INFORMATION:
; APPLICANT: M ntyl , Arja
; APPLICANT: Vehmaanpera, Jari
; APPLICANT: Fagerstrom, Richard
; APPLICANT: Lantto, Raija
; APPLICANT: Paloheimo, Marja
; APPLICANT: Suominen, Pirkko
; APPLICANT: Lahtinen, Tarja
; TITLE OF INVENTION: Production and Secretion of Proteins of
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSER: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 New York Ave., N.W. Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/770,621
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/590,563
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/332,412
; FILING DATE: 31-OCT-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/282,001
; FILING DATE: 29-JUL-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugalsky, Lawrence B.

```

```

; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 1050.0340003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1375 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: No. US20010024815A1 Relevant
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 303..1334
US-09-770-621-1

Query Match
Best Local Similarity 7.9%; Score 119.4; DB 9; Length 1375;
Matches 238; Conservative 0; Mismatches 176; Indels 3; Gaps 1;

QY 898 GTCCATAAACTACGGAGCCAACTTCCAAACCAATGGTAAATCGGTATTTATCGGTCTATGG 957
Db 608 GACCGTGACCTACAAAGCGCTCCTTCAACCCGTCGGGTAAACGGCTACCTCAGCGCTCTACGG 667

QY 958 TTGGACTGTTGACCCCTCTTGTGCAATATATATTTGTCACAGTTGGGGCAACTGCGCTCC 1017
Db 668 CTGGACCAAGAACCCGCTCGTCTGAGTACTATCATGTCGAGAGCTGGGGCACTACCGGCC 727

QY 1018 ACCAGGAGCAACGGCTAAGGGGACCATCATCTGTTGATGGAGGAACATATGATCTACGA 1077
Db 728 CACCGG---CACCTACAGGGGACCGTCACCGGCGGGGAGCGTACGACATCTACGA 784

QY 1078 GACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTCGCCACATTTAAACAATATTGGAG 1137
Db 785 GACCTGGCGGTACAAACGCGCGTCCATCGAGGGCACCCGCGACCTTCCAGCAGTTCTGGAG 844

QY 1138 TGTTTCAAGATCGAAACGACGAGTGGCAGCATTTCTGTACAGCAACCACTTTAGAGCGTG 1197
Db 845 GTCCGGCAGCAGAAAGCGGACCGGACCATCACCATCGGCAACCACTTCGACGCGCTG 904

QY 1198 GGAATACTTAGGATGAATATATGGGAAATATGAAAGTCCGCTTACTGTAGAGGCTA 1257
Db 905 GCGCCGCGCGCGCATGAACTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTA 964

QY 1258 TCAAGTAGCGGAGTGTCTATGTATATAGCAATACACTAAGAAATTAACGGTAACCC 1314
Db 965 CCAGAGCAGCGGTAGTCTCCACCGTCTCCATCAGCGAGGGTGGCAACCCCGGCAACCC 1021

```

```

RESULT 10
US-10-286-993-1
; Sequence 1, Application US/10286993
; Publication No. US20030148453A1
; GENERAL INFORMATION:
; APPLICANT: Mantyla, Arja
; APPLICANT: Paloheimo, Marja
; APPLICANT: Lantto, Raija
; APPLICANT: Fagerstrom, Richard
; APPLICANT: Lahtinen, Tarja
; APPLICANT: Suominen, Pirkko
; APPLICANT: Vehmaanpera, Jari
; TITLE OF INVENTION: Production and Secretion of Proteins in Filamentous
; TITLE OF INVENTION: Fungi
; FILE REFERENCE: 1716.0340004
; CURRENT APPLICATION NUMBER: US/10/286,993
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/120,804
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: PCT/FI97/00037
; PRIOR FILING DATE: 1997-01-24
; PRIOR APPLICATION NUMBER: US 08/590,563
; PRIOR FILING DATE: 1996-01-26

```

```

; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1375
; TYPE: DNA
; ORGANISM: Actinomadura flexuosa (Strain: DSM43186)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (303)...(1337)
; OTHER INFORMATION: Product= AM35 xylanase
US-10-286-993-1

Query Match          7.9%; Score 119.4; DB 15; Length 1375;
Best Local Similarity 57.1%; Pred. No. 7.1e-18;
Matches 238; Conservative 0; Mismatches 176; Indels 3; Gaps 1;

QY 898 GTCCATAAATAGGAGGCAACTTCCAAACAAATGGTAATGCGTATTTATGCGTCTATGG 957
Db 608 GACCGTGACCTACACGCGCTCTTCAACCGCTCGGGTAACGGCTACCTCAGCGCTCTACGG 667
QY 958 TTGGACTGTTGACCCCTCTTGTGCAATATTATTTGCGACAGTTGGGGCAACTGGGCTCC 1017
Db 668 CTGAGCAGGAACCCGCTCGTTCGAGTACTACATCGTCGAGAGCTGGGGCACTTACCGGCC 727
QY 1018 ACCAGGAGCAAGCCCTAAGGGGACCATCACTGTTGATGAGGAGAAACATATGATATCTACGA 1077
Db 728 CACCGG---CACCTCAAGAGGGACCGCTCACCCGACGGGGAGCGTAGACATCTACGA 784
QY 1078 GACTCTTAGAGTCAATCAACCCCTCCATTAAAGGGGATTTGCCACATTTTAAACAATATTGGAG 1137
Db 785 GACTCGCGGTACAACGCGCGCTCCATCGAGGCGACCCCGACCTTCAGCAGTTCTGGAG 844
QY 1138 TGTTCGAAGATCGAAGACGACGAGTGGCAGGATTTCTGTGACGAAACCACTTTAGACGCTG 1197
Db 845 CGTCCGCGCAGCAGAGCGGACCGACCATCACCATCGCGCAACCACTTCGACCGCTG 904
QY 1198 GGAACCTTAGGAGTGAATATGGGAAATGTATGAAGTCGCGCTTACTGTAGAAGGCTA 1257
Db 905 GGCGCGCGCGGATGAACCTGGCGGACGACGACTACCATGATGCGGACCGAGGCTA 964
QY 1258 TCAAAGTAGCGGAGTGCTAATGTATATATATACATACATAGAAATTAACGGTAAACCC 1314
Db 965 CCAGAGCAGCGGTAGCTCCACCGCTCTCCATCAGCGGGGTGGCAACCCCGGCAACCC 1021

RESULT 11
US-10-307-441-39
; Sequence 39, Application US/10307441
; Publication No. US20030166236A1
; GENERAL INFORMATION:
; APPLICANT: SUNG, Wing L.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Modified Xylanases Exhibiting Increased Thermophilicity
; TITLE OF INVENTION: and Alkalophilicity
; FILE REFERENCE: 027367-5006US
; CURRENT APPLICATION NUMBER: US/10/307,441
; PRIOR FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: PCT/CA01/00769
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/213,803
; PRIOR FILING DATE: 2000-05-31
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 596
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Trx
US-10-307-441-39

Query Match          7.6%; Score 115.2; DB 15; Length 596;
Best Local Similarity 57.7%; Pred. No. 4.7e-17;

```

```

Matches 226; Conservative 0; Mismatches 163; Indels 3; Gaps 1;

QY 893 AACATGTCCATAAACTACGAGCAACCTTCCAAACCAATGGTAATGCGTATTTATGCGTC 952
Db 191 AATAGGTGATCAACTTCTCTGGATCTTATATCGAATGGAAATTCATCTTAAGCGTC 250
QY 953 TATGTTGGAGCTGTTGACCCCTTGTGCAATATTATTTATGTCGACAGTTGGGGCAACTGG 1012
Db 251 TATGGCTGCTCTAGAAACCCCACTGATTGAAATATTACATTTGTGAAAAATTTTCGGTACCTAC 310
QY 1013 CGTCC---ACCAGGAGCAACGCCCTAAGGGGACCATCACTGTTGATGGAGGAACATATGAT 1069
Db 311 AATCGAGTACCGCGGCCACAAAATTAGCGAAGTCACTAGTAGTGGATCCGTATATGAT 370
QY 1070 ATCTACGAGACTCTTTAGAGTCAATCAACCCCTCCATTAAAGGGGATTTGCCACATTTAAACAA 1129
Db 371 ATCTACCGTACCCCAACGCGTTAATCAGCCATCGATCATTTGGAAACCGCCACCTTTTATCAG 430
QY 1130 TATTCGAGTGTTCGAAGATCGAAACGACGAGTGGCAGATTTCTGTGACGCAACCACTTT 1189
Db 431 TACTCGAGTGTTCGAGCTAATCATCGAGCTCCGGTTCGGTTAAATCTGCGAATCACTTT 490
QY 1190 AGAGCGTGGAAACTTAGGGATGATATGGGAAAAATGTATGAAGTCGCGCTTACTGTGA 1249
Db 491 AATGATCGGCGACACGAAAGGGTTAACCTAGGTACATGGAATATCAAAATCGTAGCGGTG 550
QY 1250 GAAGCGTATCAAAAGTAGCGGAAGTGTCTAATGT 1281
Db 551 GAAGCGTACTTCTCGAGTGGTTCGCGTAGTAT 582

RESULT 12
US-10-213-990-71
; Sequence 71, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Storms, Reg
; APPLICANT: Roemer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
; TITLE OF INVENTION: ENZYMES AND METHODS OF USE
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Aspergillus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(942)
US-10-213-990-71

Query Match          7.4%; Score 111.4; DB 14; Length 942;
Best Local Similarity 57.0%; Pred. No. 4.7e-16;
Matches 224; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

QY 882 AACAAAGTTGGTAACTGTCCATAAACTACGAGGCCAACTTCCAAACCAATGGTAATGCGT 941
Db 257 ATCCAGGGAGTGACCATGACATTAACCTTCTCTGGCAGCTTCAATCTCTTCGGAATGCTT 316
QY 942 ATTTATGCTCTATGTTGGACCTGTTGACCTCTTCTGCAATATTATTTATGTCGACAGTT 1001
Db 317 ACCTGTCCGTGTATGATGGACTACCAACCCCTTAGTCGAATACTCATCTCTCGAAGACT 376
QY 1002 GGGGCA---ACTCGCGTCCACAGGAGCAACGCGTAAAGGGGACCATCACTGTTGATGGAG 1058
Db 377 ATGGCAGTTTAAATCTCTGGCTCGGCGATGACGACAGGGCACCGTCACCGAGATGAT 436
QY 1059 GAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATGGCA 1118

```

Db 437 CCACCTACGACATCTATGAGCAACCAACAGGTCAACAGCCTTCGATCGTCGCGCACGCCA 496  
Qy 1119 CATTTAAACAATATGAGTGTTCGAGATCGAAACGACGAGTGGCAGCATTTCTGTCA 1178  
Db 497 CCTTCAACCAATCTAGTCCATCCGCGCAAAACAAGCATCCAGCGGCACAGTCACCAACG 556  
Qy 1179 GAACACACTTTAGAGCGTGGGAAACTTAGGATGAATATGGGAAATGTATGAGTCG 1238  
Db 557 CGAATCACHTCAGGCGCTGGCTAGTCTGGGATGAACCTGGTACCCATACTATCAGA 616  
Qy 1239 CGCTTACTAGAGGCTATCAAAAGTAGCGGAA 1271  
Db 617 TTGTTTCCACTGAGGATATGAGACGCGGTA 649

## RESULT 13

US-10-213-990-70  
; Sequence 70, Application US/10213990  
; Publication No. US20030082595A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Bussey, Howard  
; APPLICANT: Roemer, Terry  
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL  
; FILE REFERENCE: 10182-019-999  
; CURRENT APPLICATION NUMBER: US/10/213,990  
; CURRENT FILING DATE: 2002-08-05  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 70  
; LENGTH: 1002  
; TYPE: DNA  
; ORGANISM: Aspergillus  
US-10-213-990-70

Query Match 7.3%; Score 110; DB 14; Length 1002;

Best Local Similarity 57.8%; Pred. No. 1e-15; Indels 3; Gaps 1;  
Matches 216; Conservative 0; Mismatches 155;

Qy 901 CATAACTACGAGGCAACTTCCAAACCAATGGTAATGCGTATTTATGCGTCTATGGTTG 960  
Db 336 CATTACCTTCTGCGCAGCTTCAATCCTTCGGGAAATGCTTACCTGTCGCTATGGATG 395  
Qy 961 GACTGTGACCTCTTGTGCAATATATATGTCACAGTTGGGCA---ACTGCGCTCC 1017  
Db 396 GACTACCAACCCCTTAGTGAATATACATCTCGAGAACTATGCGAGTTACAATCCTGG 455  
Qy 1018 ACCAGGACCAACGCTTAAGGGGACCATCACTGTTGATGGAGAACATATGATCTACGA 1077  
Db 456 CTCGGGCATGCGACAGAGGACCGTACACAGCGATGATCCACTAGACATCTATGA 515  
Qy 1078 GACTCTTAGAGTCAATCAACCTCCATTAAGGGGATGCCATTTAAACAATATPGAG 1137  
Db 516 GCACCAACAGGTCAACAGCCTTCGATCGTCGCGCACGCACTTCAACCAATCTGCTC 575  
Qy 1138 TGTTTGAAGATGAAACGACGAGTGGCAGATTTCTGTACGACCACTTTAGAGCGTG 1197  
Db 576 CATCGGCCAAAACAAGCGATCCAGCGGCACAGTCAACCGCGAATCACTTCAAGGCTG 635  
Qy 1198 GGAACACTTAGGATGAATATGGGAAATGTATGAGTCGGGCTTACTGTAGAGGCTA 1257  
Db 636 GCGTAGTCTGGGATGAACCTGGGTACCCATACTATCAGATTGTTTCCACTGAGGATA 695  
Qy 1258 TCAAGTAGCGGAA 1271  
Db 696 TGAGACGACGGTA 709

## RESULT 14

US-10-213-990-68

; Sequence 68, Application US/10213990  
; Publication No. US20030082595A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Bussey, Howard  
; APPLICANT: Roemer, Terry  
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL  
; FILE REFERENCE: 10182-019-999  
; CURRENT APPLICATION NUMBER: US/10/213,990  
; CURRENT FILING DATE: 2002-08-05  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 68  
; LENGTH: 705  
; TYPE: DNA  
; ORGANISM: Aspergillus  
; NAME/KEY: CDS  
; LOCATION: (1)...(705)  
US-10-213-990-68

Query Match 7.2%; Score 108.8; DB 14; Length 705;

Best Local Similarity 51.9%; Pred. No. 1.7e-15;  
Matches 298; Conservative 0; Mismatches 267; Indels 9; Gaps 2;

Qy 706 CGTCACGACAAATTCCTTGGCAACCAACGATGGCTATGATTAATTTTGGAAAGATAG 765  
Db 120 CGGACACCCCAAGCTCCACCGCTGGAAACAACGGGTACTACTCTCTTGGACTGTGG 179  
Qy 766 CGGTGGCTCTGGGACAATGATTCATCATCTGCGGTGCTGCTGCTGCTGCTGCTGCTG 825  
Db 180 CGCGCGCACTGTGACCTACCAATGGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTG 239  
Qy 826 TGTTAAACAACATATTTATTCGTTAAAGGTAAATAATTTCAATGAACACACAAACACCA 885  
Db 240 CGTGGCAACTTTGTCGTGGGAAAGGCG-----TGAACCCCTGGAAGCGCTAGGTACCG 293  
Qy 886 AGTTGGTAACATGCTTAACATACGAGGACCACTTCCAACCAATGGTATGCGTATTT 945  
Db 294 AGCTTAAAGTAGAACCACTCACTACGAGGCGAGCTTCAACCCCGAGCGCAATGGCTACCT 353  
Qy 946 ATGCGTCTATGGTTGGACTGTTGACCTCTTGTGCAATATTTATTTGCGACAGTTGGGG 1005  
Db 354 GGCTGTCTACGCTGGACCAACCCCTTGATTGAGTACTACCTGTTGAGTCGTATGG 413  
Qy 1006 CAATCGCGTCCACA---GGAGCAACGCTTAAGGGGACCATCACTGTTGATGAGGAAC 1062  
Db 414 TACATACAAACCCCGCGAGCGCGGTACTCTCAGGGGCACTGTCAACACCGACGCTGGCAC 473  
Qy 1063 ATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATGGCCACATT 1122  
Db 474 TTAACAATCTACACGCGCTTCGTACATGCTCCCTCCATCGAAGGCACCAAGACCTT 533  
Qy 1123 TAAACAATATTTGAGTGTTCGAAGATGAAACGACGAGTGGGACAGATTTCTGTGAGCAA 1182  
Db 534 CACCCAGTACTGGTCTGTGCGCACTTCAAGCGTACCGGCGCACTGTCAACATGGGCAA 593  
Qy 1183 CCATTTAGAGCGTGGGAAACCTTAGGATGAATATGGGAAATGTATGAGTGGCGCT 1242  
Db 594 CCATTTCAACCCCTTGGAGCAGACTGGGCGATGAACCTGGGAACCTCACAACCTACAGATTGT 653  
Qy 1243 TACTGTAGAGGCTATCAAGTAGTACGGAAGTGTCT 1276  
Db 654 CGCCACTGAGGTTTACCAGACGCGGATCTGCT 687

## RESULT 15

US-10-299-393-1  
; Sequence 1, Application US/10299393  
; Publication No. US20030108642A1  
; GENERAL INFORMATION:



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 16:24:32 ; Search time 6558.71 Seconds  
(without alignments)  
8406.125 Million cell updates/sec

Title: US-09-909-207-10  
Perfect score: 1513  
Sequence: 1 AAATTGATTGGTATATCT.....TGAAACACCTCGTCACTAG 1513

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98.6	6.5	618	7	CF472462 RTDS1_9 C
2	93	6.1	728	7	CN132941 OX1_9_D10
3	93	6.1	772	7	CN133022 OX1_9_D10
4	91.8	6.1	744	7	CF867983 trico13xe
5	91.8	6.1	799	6	CB898036 trico13xe
6	86.6	5.7	738	6	CD464145 ETH1_48 B
7	83.4	5.5	921	6	CD458837 Fg08_04B1
8	82.8	5.5	603	8	AQ160513 mgxb0006C
9	82.2	5.4	670	8	AQ447125 mgxb0005C
10	82.2	5.4	720	8	AQ361561 mgxb0004B
11	82.2	5.4	750	8	AQ160254 mgxb0003L
12	81.6	5.4	786	8	AQ325248 mgxb00021M
13	81.4	5.4	583	8	AQ399120 mgxb0001B
14	79.8	5.3	1101	9	AL057419 Drosophil
15	76.6	5.1	561	8	AQ396475 mgxb0010M
16	76.4	5.0	617	7	AJ638869 AJ638869
17	76.2	5.0	768	7	CF881056 trico083xj
18	76.2	5.0	822	6	CB907827 trico083xj
19	74.6	4.9	753	8	AQ448084 mgxb0016B
20	74.4	4.9	418	8	AQ398756 mgxb0005L
21	71.8	4.7	746	6	CB901964 trico028xi
22	71.8	4.7	746	7	CF871731 trico028xi
23	71.4	4.7	1101	9	AL063921 Drosophil
24	69.8	4.6	520	5	BQ165950 WHE0821-0

c	25	67.6	4.5	1101	9	CNS00EVL	AL069706 Drosophil
	26	66.2	4.4	561	5	BQ471960	CD471960 HV04A02r
	27	63.2	4.2	473	6	CD464005	ETH1_48 B
c	28	62	4.1	1101	9	CNS0021J	AL061936 Drosophil
c	29	61	4.0	939	6	BY720774	BY720774 BY720774
	30	60.6	4.0	1101	9	CNS00EVL	AL069706 Drosophil
c	31	60.4	4.0	483	5	BQ664593	BQ664593 HV04A02u
c	32	60.2	4.0	1068	8	CC263147	CC263147 CH261-168
c	33	58.4	3.9	1101	9	CNS000D1	AL065414 Drosophil
	34	57.2	3.8	1147	6	CD387645	CD387645 AGENCOURT
	35	56.4	3.7	1031	9	CNS00CF2	AL059199 Drosophil
c	36	56	3.7	1001	9	CNS01400	AL103554 Drosophil
	37	54.8	3.6	886	8	BH177277	BH177277 008_L_22-
	38	54.8	3.6	886	9	CNS07JUX	AL061423 T3 end of
	39	54.8	3.6	1101	9	CNS00BO7	AL069440 Drosophil
	40	54.6	3.6	843	9	CNS00CS1	AL059666 Drosophil
	41	54.6	3.6	958	9	CNS04BNX	AL304134 Tetraodon
c	42	54.4	3.6	1092	9	CNS020K7	AL175696 Tetraodon
	43	54.2	3.6	1043	9	CNS0145P	AL103735 Drosophil
	44	54	3.6	928	9	CNS00DKY	AL071865 Drosophil
c	45	54	3.6	928	9	CNS00DKY	AL071865 Drosophil

ALIGNMENTS

RESULT 1  
CF472462  
LOCUS  
DEFINITION  
RTDS1\_9 C03.g1\_A015 Drought-stressed loblolly pine roots DS1 Pinus taeda CDS clone RTDS1\_9\_C03\_A015 5', mRNA sequence.  
CF472462  
CF472462.1 GI:34489834  
EST.  
Pinus taeda (loblolly pine)  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J., Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and Neale, D.  
TITLE  
An EST database from drought-stressed loblolly pine (Pinus taeda) roots  
JOURNAL  
COMMENT  
Unpublished (2003)  
Other ESTs: RTDS1\_9 C03.b1\_A015  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@cuga.edu  
RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
Seq primer: JENREV (CAGGAACAGCTATGACC).  
Location/Qualifiers  
1. .618  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="CCLONES"  
/db\_xref="taxon:3352"  
/clone="RTDS1\_9 C03 A015"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="Drought-stressed loblolly pine roots DS1"  
/notes="Vector: pSL180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from drought-stressed loblolly pine (Pinus taeda) roots. Water was withheld from

ramet clones until predawn needle water potential reached -1.75 MPa. On day 7 roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSul180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

# ORIGIN

Query Match 6.5%; Score 98.6; DB 7; Length 618;  
 Best Local Similarity 57.4%; Pred. No. 2.8e-12;  
 Matches 198; Conservative 0; Mismatches 144; Indels 3; Gaps 1;  
 QY 882 AACAAAGTGGTAACATGTCCATAAAGTACGAGCAGCAACTTCCAAACCAATGTAATGCT 941  
 Db 254 ATCCAGGTAGTCGCGAGGACATCACTTCTCGGTACCTTCAGCCCGCAGAAAGCT 313  
 QY 942 ATTATGCGTCTATGGTGGAGCTGTGACCCCTTGTGCAATATATATATATGTCGACGTT 1001  
 Db 314 ATCTTGGCGTCTACGGCTGGACCAAGCCCTCTGATCGAATATACATTTCTCGAAAGCT 373  
 QY 1002 GGGCAACTGGGCTCC---ACCAGAGCAAGCCCTTAAGGGACCATCACTGTTGATGGAG 1058  
 Db 374 ATGGGACCTTCAACCCCTGGATCCAGATGACCATATGGGCACTGTCCAGCGAGCTG 433  
 QY 1059 GAACATATGATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAAGGGGATTGCCA 1118  
 Db 434 CGACCTATGATCTATGAGCAGCAGCAAGTCAACAGCCCTCTATCTGCGCAGCGGA 493  
 QY 1119 CATTAAACAATATGAGTGTGCAAGATCGAAGCAAGCAAGTGGCGAGATTCTGTCA 1178  
 Db 494 CTTTCAACCAATATCTGCTCCATCTCGTCAGAGCAAGCGGTCCAGCGGAATGTCACTACAG 553  
 QY 1179 GCAACCACTTATAGCGTGGGAAACTTAGGATCAATATGGGA 1223  
 Db 554 CGAACCAATTTCAACGCTGGGCTGCTTGGGCATGAACCTGGGGA 598

# RESULT 2

CN132941  
 LOCUS OX1\_9\_D10.bl.A002 728 bp mRNA linear EST 01-APR-2004  
 DEFINITION bicolor cDNA clone OX1\_9\_D10\_A002 3', mRNA sequence.  
 ACCESSION CN132941  
 VERSION CN132941.1 GI:45963408  
 KEYWORDS EST.  
 SOURCE Sorghum bicolor (sorghum)  
 ORGANISM Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.  
 1 (bases 1 to 728)  
 Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S., Klein,R.R., Liang,C., Sun,F., Sullivan,R., Lim,S., Eastman,A. and Pratt,L.H.  
 An EST database from Sorghum: oxidatively stressed leaves and roots Unpublished (2003)  
 Other ESTs: OX1\_9\_D10.g1.A002  
 Contact: Cordonnier-Pratt MW  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu  
 Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University; sequencing done in the laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
 Seq primer: Sug3-14 (TACTTACGGCGCGGACCC)  
 POLYA=Yes.

# FEATURES

Location/Qualifiers

# source

1. 728  
 /organism="Sorghum bicolor"  
 /molt\_type="mRNA"  
 /db\_xref="taxon:4558"  
 /clones="OX1\_9\_D10\_A002"  
 /lab\_host="DH10B-T1 phage-resistant E. coli"  
 /clone\_lib="Oxidatively-stressed leaves and roots"  
 /note="Organ: Leaf and Root; Vector: pME18S-FL3; Site.1: RNA from oxidatively stressed, hydroponically grown sorghum seedlings. At 8 days of age, growth medium was supplemented with hydrogen peroxide to 0.003% and leaves were misted with 10 uM methyl viologen. Leaves and roots were harvested at 3, 12 and 27 hr after treatment and all tissue pooled. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."

# ORIGIN

Query Match 6.1%; Score 93; DB 7; Length 728;  
 Best Local Similarity 53.9%; Pred. No. 5.8e-11;  
 Matches 214; Conservative 0; Mismatches 180; Indels 3; Gaps 1;  
 QY 900 CCATAAAGTACGAGCGCAACTTCCAAACCAATGTAATGCTATTTATGCTCTATGTT 959  
 Db 101 CCATCAACTATGGCGGTTCTTTTCAGCCCTCAGGCTAAGGCTCTCTGCGTCTACGGCT 160  
 QY 960 GGACTGTTGACCCCTCTTGTGCAATATATATGTCGACAGTTGGGCA---ACTGGCGTC 1016  
 Db 161 GGACTCGAGCCCTCTGTTGAGTACTACGTCATCGAGAACTACGGCACTTACACCTG 220  
 QY 1017 CACCAGGAGCAACGCCCTTAAGGGGACCATCACTGTTGATGGAGAACATATATATCTACG 1076  
 Db 221 GCTCTGCTGGCCAGCACAAGGGGCAACCGTCTACAACGCGGACACCTTACGATCTCTACC 280  
 QY 1077 AGACTCTTAGAGTCAATCAACCCCTCCATTAGGGGATGCCACATTTAAACAATATGGA 1136  
 Db 281 AGACCCACCCGCTACAACCCAGCCCTCTATCGACGCGCAACAGACCTTCAACAGTACTGGG 340  
 QY 1137 GTGTTGGAAGATCGAAACGACGAGTGCAGATTTCTGTGACGAACCACTTTTAGAGCGT 1196  
 Db 341 CCATCCCGCCCAACAGCGCAGCGCGCGCTCAACATCGACATCTTCAATGTT 400  
 QY 1197 GGGAAAATTTAGGATGAATATGGGAAAATGTTATGAAGTCGCGCTTACTGTAGAGGCT 1256  
 Db 401 GGGCTAACGCTGTTATGAGACTTGGAAACCACTACTACAGATCTGCTGCTACCGAGGAT 460  
 QY 1257 ATCAAGTAGCGGAAGTCTTAATGTTATATAGCAATAC 1293  
 Db 461 ACCAGAGCAGTGGATCTTCTTCTATCTATGTCAGAC 497

# RESULT 3

CN133022  
 LOCUS OX1\_9\_D10.g1.A002 772 bp mRNA linear EST 01-APR-2004  
 DEFINITION bicolor cDNA clone OX1\_9\_D10\_A002 5', mRNA sequence.  
 ACCESSION CN133022  
 VERSION CN133022.1 GI:45963542  
 KEYWORDS EST.  
 SOURCE Sorghum bicolor (sorghum)  
 ORGANISM Sorghum bicolor  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.  
 1 (bases 1 to 772)  
 Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S., Klein,R.R., Liang,C., Sun,F., Sullivan,R., Lim,S., Eastman,A. and Pratt,L.H.  
 An EST database from Sorghum: oxidatively stressed leaves and roots Unpublished (2003)



## COMMENT

Other ESTs: OX1.9.D10.b1.A002  
 Contact: Cordomier-Pratt NM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu  
 Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in  
 the Human Genome Center, University of Tokyo Institute of Medical  
 Science; plant material and RNA prepared at Texas A & M University;  
 sequencing done in the Laboratory for Genomics and Bioinformatics,  
 University of Georgia. Sequence ends have been trimmed to exclude  
 vector and regions below Phred quality 16. Three-prime sequences  
 are presented as their reverse complement and have been trimmed to  
 exclude polyA.  
 Seq primer: Sug5 (CTTGTGCTTAAAGTGGC).

## FEATURES

Source

1. .772  
 Location/Qualifiers  
 /organism="Sorghum bicolor"  
 /mol\_type="mRNA"  
 /cultivar="BTx623"  
 /db\_xref="taxon:4558"  
 /clone="OX1.9.D10.A002"  
 /lab\_host="DH10B-Ti phage-resistant E. coli"  
 /clone\_lib="Oxidatively-stressed leaves and roots"  
 /note="Organ: Leaf and Root; Vector: pHE18S-FL3; Site: 1:  
 XhoI; Site: 2: XhoI; The library was prepared from polyA+  
 RNA from oxidatively stressed, hydroponically grown  
 sorghum seedlings. At 8 days of age, growth medium was  
 supplemented with hydrogen peroxide to 0.003% and leaves  
 were misted with 10 uM methyl viologen. Leaves and roots  
 were harvested at 3, 12 and 27 hr after treatment and all  
 tissue pooled. Double-stranded cDNA was cloned  
 unidirectionally into different DraIII sites of the  
 pHE18S-FL3 vector (5-prime DraIII site is CACTGTGTG,  
 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA  
 insert."

## ORIGIN

Query Match 6.1%; Score 93; DB 7; Length 772;  
 Best Local Similarity 53.3%; Pred. No. 5.9e-11;  
 Matches 214; Conservative 0; Mismatches 180; Indels 3; Gaps 1;  
 QY 900 CCATAAAGTACGAGGCACTTCCAAACCAATGGTAATGCGTATTTATGCGTCTATGGTT 959  
 Db |||||  
 324 CCATCAACTATGCGGGTTCTTTCAGCCCTCAGGGTAACGGCTACCTCTGCGTCTACGGCT 383  
 QY |||||  
 960 GGACTGTGACCTCTTGTGCAATATATATTCTGCACAGTTGGGCA---ACTGGCGTC 1016  
 Db |||||  
 384 GGACTGCGACCTCTCGTTGGTACTACTGCTATCGAGAACTACGGCACTTACACCTG 443  
 QY |||||  
 1017 CACCAGGAGCAACGCTTACGAGGACCATCACTGTGATGGAGGAACATATGATCTACG 1076  
 Db |||||  
 444 GCTCTGCTGCCAGCACAAGGCGACCGTCTACAACGAGCGGACACTACGACTCTTACC 503  
 QY |||||  
 1077 AGACTCTTAGAGTCAATCAACCTTCATTAAGGGGATGGCCATTAAACAATATTGA 1136  
 Db |||||  
 504 AGACCAACCCCTACAACAGCCCTCTATCGACGGGCCAACAGACCTTCAACCACTACTGGG 563  
 QY |||||  
 1137 GTCTTGAAGATCGAAGCAGGAGTGCGACGATTTCTGTGAGCAACCACTTTTAGAGCGT 1196  
 Db |||||  
 564 CCATCCGCCGCAACAAGCGAGAGCGCGCGCTCAACATGCGAGACTATCTTCATGCTT 623  
 QY |||||  
 1197 GGGAAACTTAGGGATGAATATGGGAAAAATGTATGAAGTTCGGCTTACTGTAGAGGCT 1256  
 Db |||||  
 624 GGGCTAACGCTGGTATGAGACTTGGAAACCACTACTACAGATCCTTGGCTACCGAGGAT 683  
 QY |||||  
 1257 ATCAAGTACGGAAGTGCCTAAATGTATATAGCAATAC 1293  
 Db |||||  
 684 ACCAGACAGTGGATCTTCTTCTATCTATCTATGTCCAGAC 720

## RESULT 4

CF867983 744 bp mRNA linear EST 31-OCT-2003  
 LOCUS tricol13xe09.b1 T.reesei mycelial culture, Version 6 October 2003  
 DEFINITION Hypocrea jecorina cDNA clone tricol13xe09, mRNA sequence.  
 ACCESSION CF867983  
 VERSION 1  
 KEYWORDS EST.  
 SOURCE Hypocrea jecorina (anamorph: Trichoderma reesei)  
 ORGANISM Hypocrea jecorina  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.  
 REFERENCE 1 (bases 1 to 744)  
 AUTHORS Diener,S.E., Dunn-Coleman,N., Foreman,P., Houfek,T.D.,  
 Teunissen,P.J.M., van Solingen,P., Dankmeyer,L., Mitchell,T.K.,  
 Ward,M. and Dean,R.A.  
 TITLE Characterization of the protein processing and secretion pathways  
 in a comprehensive set of expressed sequence tags from Trichoderma  
 reesei  
 JOURNAL PEMS Microbiol. Lett. 230 (2), 275-282 (2004)  
 COMMENT Contact: Ralph A. Dean  
 Fungal Genomics Laboratory  
 North Carolina State University  
 Campus Box 7251, Raleigh, NC 27695, USA  
 Tel: 919-513-0020  
 Fax: 919-513-0024  
 Email: ralph.dean@ncsu.edu  
 Seq primer: Lr-F1 primer.  
 FEATURES  
 source  
 1. .744  
 /organism="Hypocrea jecorina"  
 /mol\_type="mRNA"  
 /strain="QM6a"  
 /db\_xref="taxon:51453"  
 /clone="tricol13xe09"  
 /dev\_stage="mycelia"  
 /clone\_lib="T.reesei mycelial culture, Version 6 October  
 2003"  
 /note="Vector: pRBP3Y; Site 1: Not I/Sal I; Mycelial  
 culture grown from 24 hrs to 6 days with varying Carbon  
 and Nitrogen sources and concentrations."  
 ORIGIN  
 Query Match 6.1%; Score 91.8; DB 7; Length 744;  
 Best Local Similarity 55.0%; Pred. No. 1.1e-10;  
 Matches 202; Conservative 0; Mismatches 162; Indels 3; Gaps 1;  
 QY 893 AACATGTCCATAAACTACGGAGCCAACTTCCAAACCAATGGTAATGCGTATTTATGCGTC 952  
 Db |||||  
 353 AACNAGGTCATCAACTTCTCGGGCAGCTACACCCCAACGGCAACAGCTACCTCTCCGTG 412  
 QY |||||  
 953 TATGGTTGGACTGTTGACCCCTCTTGTGCAATATATATTGTCGACAGTTGGGGCAACTGG 1012  
 Db |||||  
 413 TAGCGGTGTCCCGCAACCCCTGTATGAGTACTATCATGTCGAGAACTTTGGCACTAC 472  
 QY |||||  
 1013 CGTCCACCA---GGAGCAACGCCAAGGGGACCATCACTGTTGATGGAGGAACATATGAT 1069  
 Db |||||  
 473 AACCCGTCCACGGGCGCCACCAAGCTGGCGAGGTCACTCCGACGGCAGCGCTACGAC 532  
 QY |||||  
 1070 ATCTACGAGACTTTAGAGTCAATCAACCTCCATTAAGGGGATTCGCAATTAAACAA 1129  
 Db |||||  
 533 ATTTACCGCAGCAGCGCGTCAACCAAGCGCTCCATCATCGGCAACCGCCACTTTTACCAG 592  
 QY |||||  
 1130 TATTGAGTGTTCGAGATCGAACCACGACGAGTGGCACGATTTCTGTACGACCAACCTTT 1189  
 Db |||||  
 593 TACTGTCTCGTCGCGCGCAACCCGCTGAGGGGTCCGTCACACGGCGGAACCACTTC 652  
 QY |||||  
 1190 AGAGCGTGGGAAACTTAGGGATGAATATGGGAAAAATGTATCAAGTCGCGCTTACTGTGA 1249  
 Db |||||  
 653 AACCGTGGGCTCAGCAAGGCCTGACGCTCGGACGATGATACAGATTGTTGCCGTG 712  
 QY |||||  
 1250 GAAGGCT 1256

```

Db          713 GGAGGGT 719

RESULT 5
CB898036
LOCUS
DEFINITION
trio13xe09 T.reesei mycelial culture, Version 3 april Hypocrea
jecorina cDNA clone trio13xe09, mRNA sequence.
ACCESSION
CB898036
VERSION
CB898036.1 GI:30112694
KEYWORDS
EST.
SOURCE
Hypocrea jecorina (anamorph: Trichoderma reesei)
ORGANISM
Hypocrea jecorina
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
REFERENCE
Foreman,P.K., Brown,D.E., Dankmeyer,L., Dean,R., Diener,S.,
Dunn-Coleman,N.S., Goedegebuur,F., Houfek,T.D., England,G.J.,
Kelley,A.S., Meerman,H.J., Mitchell,T., Mitchinson,C.,
Olivares,H.A., Teunissen,P.J., Yao,J. and Ward,M.
Transcriptional regulation of biomass-degrading enzymes in the
filamentous fungus Trichoderma reesei
J. Biol. Chem. 278 (34), 31988-31997 (2003)
22803314
PUBMED
12788920
COMMENT
Contact: Pamela K. Foreman
Genencor Intl.
925 Page Mill Road, Palo Alto, CA 94304, USA
Tel: (650) 846-7635
Fax: (650) 621-7817
Email: pforeman@genencor.com
Seq primer: LT-Fl primer.
FEATURES
source
1..799
/organism="Hypocrea jecorina"
/mol_type="mRNA"
/strain="QM6a"
/db_xref="taxon:51453"
/dev_stage="mycelia"
/clone_lib="trio13xe09"
/clone_lib="T.reesei mycelial culture, Version 3 april"
/vector="pREP3v; Site 1: Not I/Sal I; Mycelial
culture grown from 24 hrs to 6 days with varying Carbon
and Nitrogen sources and concentrations."
ORIGIN
Query Match 6.1%; Score 91.8; DB 6; Length 799;
Best Local Similarity 55.0%; Pred. No. 1.1e-10;
Matches 202; Conservative 0; Mismatches 162; Indels 3; Gaps 1;
QY 893 AACATGTCATTAAGTACGGAGCCCACTTCCAAACCAATGTAATGCGTATTATGCGTC 952
Db 408 AACAAAGGTATCACTTCTCGGGCAGCTACACCCCAACGGCAACAGTACTCTCCGTG 467
QY 953 TATGTTGGACCTGTTGACCTCTTGTGCAATATTATTATGTCGACAGTTGGGCAACTGG 1012
Db 468 TAGCGTGTCCGGCAACCCCTGATCGAGTACTACATCGTCGAGAACTTGGACCTAC 527
QY 1013 CGTCCACCA---GGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGAACATATGAT 1069
Db 528 AACCCGTCCACGGGCGCCACCAAGCTGGGCGAGGTCACTCCGACGGCAGCGTCTACGAC 587
QY 1070 ATCTAGGAGACTTATAGTGTCAATCAACCTCCATTAAGGGGATGCCACATTTAAACAA 1129
Db 588 ATTTACCGCAGCAGCGGTCAACAGCGGTCCATCATCGGACCGCCACCTTTTACCAG 647
QY 1130 TATTGAGTGTTCGAAGATCGAAACCGCAGAGTGCAGGATTTCTGTCAGCAACACACTTT 1189
Db 648 TACTGTCCGTCCGGCGCAACACCGCTCGAGCGCTCCGTCAACACGGCGNACCACTTC 707
QY 1190 AGAGCGTGGGAAACATTAGGATGATATGGGAAATATGATGAAGTCGCGTACTCTGTA 1249
Db 708 AACCGTGGGCTCAGCAAGGCGCTGCGCTCGGACGATGGATTACAGATTGTTGCCGTG 767

QY 1250 GAAGGCT 1256
Db 768 GGAGGGT 774

RESULT 6
CD464145
LOCUS
DEFINITION
ETH1_48_B06_g1_A002 Ethylene-treated seedlings Sorghum bicolor cDNA
clone ETH1_48_B06_A002 5', mRNA sequence.
ACCESSION
CD464145
VERSION
CD464145.1 GI:31385413
KEYWORDS
EST.
SOURCE
Sorghum bicolor (sorghum)
ORGANISM
Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE
1 (bases 1 to 738)
Cordonnier-Pratt,M.-M., Wentzel,V., Suzuki,Y., Sugano,S.,
Klein,R.R., Liang,C., Sun,F., Sullivan,R., Shah,M., Salzman,R.,
Chua Tan,N., Gonzalez,M., Lane,S., Miller,V., Nanda,P.,
Olaseinde,O., Eastman,A. and Pratt,L.H.
An EST database from Sorghum: 1-aminocyclopropane-1-carboxylic acid
(ACC)-treated seedlings
Unpublished (2003)
Other ESTs: ETH1_48_B06_b1_A002
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq primer: Sug5 (CTTCGCTCTAAAGCTGGC).
FEATURES
source
1..738
/organism="Sorghum bicolor"
/mol_type="mRNA"
/cultivar="BTx623"
/db_xref="taxon:4558"
/clone="ETH1_48_B06_A002"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="Ethylene-treated seedlings"
/vector="pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The
library was prepared from polyA+ RNA from seedlings grown
in hydroponic culture. At 8 days of age, medium was
supplemented with 0.5 mM 1-aminocyclopropane-1-carboxylic
acid (ACC) to induce endogenous ethylene (ETH) production.
Roots and shoots were harvested after 27 and 72 hr and
material from both time points was combined prior to RNA
isolation. Double-stranded cDNA was cloned
unidirectionally into different DraIII sites of the
pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG,
3-prime DraIII site is CACCATGTG). XhoI excises the cDNA
insert."
ORIGIN
Query Match 5.7%; Score 86.6; DB 6; Length 738;
Best Local Similarity 53.5%; Pred. No. 1.8e-09;
Matches 204; Conservative 0; Mismatches 174; Indels 3; Gaps 1;
QY 900 CCATAAATACGGAGCCCACTTCCAAACCAATGTAATGCGTATTATGCGTCTATGTT 959

```

Db 345 CCATCAAGTACTCTGTCTTACACATCAACGAAACAGCTACCTCGCTGTTTACGGAT 404

QY 960 GGACTGTTGACCTCTTGTGCAATATATATTGTGACAGTTGGGGCACTGGCTCCAC 1019

Db 405 GGACTCAGAACCTCTCTCATCGAGTACTACATCGTTGAGAACTTCGGCACCTTCAACCCCT 464

QY 1020 CA--GGAGCAACGCTAAGGGGACCATCACTGTTGATGGAGAACATATATATCTAGC 1076

Db 465 CTTCCGGCGCCGAGAGAGGTTGAGGTCACTGTTGACGGATCTGTCTACGACATCTAGC 524

QY 1077 AGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTCACATTTAAACAATATGGA 1136

Db 525 TCAGCACCGGTGTCAACGCCCCCTCATTTGAGGTAACAAGACCTTTTACAGCACTTCTGT 584

QY 1137 GTGTTGCAAGATCGAAGCGACGAGTGGCAGGATTTCTGTGACCAACCACTTTAGAGCGT 1196

Db 585 CTGTTCCGACCAACAGCGATCCAGCGGATCCGTCACACCGGTGCTCACTTCCAGGCT 644

QY 1197 GGGAAACTTAGGATGAATATGGGAAATGTATGAAGTCGCGCTTACTGTAGAGGCT 1256

Db 645 GGGAAATGTGCGGCTCAACCTTGGTAAACCAACTATCAGATCCTTGTGTTGAGGCT 704

QY 1257 ATCAAGTAGCGGAAGTGCTA 1277

Db 705 ACTACAGCTCGGCTGTGCCA 725

RESULT 7

LOCUS CD458837 921 bp mRNA linear EST 14-JUN-2004

DEFINITION Fg08\_04b10. A Fg08\_AAPC\_ECORC\_Fusarium graminearum complex\_substrate

ACCESSION CD458837

VERSION CD458837.1 GI:31373577

KEYWORDS EST.

SOURCE Gibberella zeae

ORGANISM Gibberella zeae

REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.

AUTHORS Watson, R.J., Heyes, R., Chapados, J., Couroux, P., Harris, L.J., Hattori, J., Lacroix, C., Ouellet, T., Robert, L.S., Singh, J.A., Spratt, D. and Tinker, N.A.

TITLE A cDNA library prepared from Fusarium graminearum grown on a complex plant substrate

JOURNAL Unpublished (2003)

COMMENT Contact: Watson, Robert.J.  
Eastern Cereal and Oilseed Research Centre  
Agriculture and Agri-Food Canada  
Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6, CANADA

FEATURES

source

1. .921

/organism="Gibberella zeae"

/mol\_type="mRNA"

/strain="DAOM 180378"

/db\_xref="taxon:5518"

/clone="Fg08\_04b10"

/tissue\_type="Mycelium"

/dev\_stage="Asexual"

/lab\_host="E. coli DH10B"

/clone\_lib="Fg08\_AAPC\_ECORC\_Fusarium graminearum\_complex\_substrate"

/note="Vector: pBluescript II+; Site 1: EcoRI; Site 2: XhoI; Fusarium graminearum grown on a complex plant substrate-- wheat leaves treated to remove most of the low molecular weight, water-soluble components."

ORIGIN

Query Match

5.5%; Score 83.4; DB 6; Length 921;

Best Local Similarity 52.4%; Pred. No. 1.1e-08;  
Matches 208; Conservative 0; Mismatches 186; Indels 3; Gaps 1;

QY 900 CCATCAAGTACTCGGAGCAACTTCCAAACCAATGTAATGCGTATTTATGCGCTATGGTT 959

Db 363 CCATCAAGTACTCGGAGGTTCCCTTCAACCTCAGGGTAACGGATACCTTTCGGTTACGGAT 422

QY 960 GGACTGTTGACCTCTTGTGCAATATATATTGTGCGAGTTGGGGCACTGGCGTCCAC 1019

Db 423 GGAACCGCGGTCCCTCGTCGAGTACTAGCTCATCGAGATTACGGTTCTTCAACACCCG 482

QY 1020 CAGGA--GCAACGCTAAGGGGACCATCACTGTTGATGGAGAACATATATATCTAGC 1076

Db 483 CGAGCCAGGCTCAGACCGAGGTACCGTCTACACGACGGTGACACCTACGATCTCTATA 542

QY 1077 AGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTCGACATTTAAACAATATGGA 1136

Db 543 TGTCCACCGGTTACCAACAGCCTTCATCGACGGTGTTCAGACCTTCAACAGTACTGGT 602

QY 1137 GTGTTGCAAGATCGAAGCGACGAGTGGCAGCATTTCTGTGACCAACCACTTTAGAGCGT 1196

Db 603 CCATCGCGCGCAACAGCGTACACGGGCTCGTCAACATGCGAAGCACTTCAATGCTT 662

QY 1197 GGGAAACTTAGGATGAATATGGGAAATGTATGAAGTCGCGCTTACTGTAGAGGCT 1256

Db 663 GGAGATCTGCTGGCATGAACCTCGGAAACCACTACTACCAGATTCTGGCCACTGAGGGTT 722

QY 1257 ATCAAGTAGCGGAAGTGCTAATGTATATAGCAATAC 1293

Db 723 ACCAGAGCAGTGGCTCATCTTCTATCTATGTCCAGAC 759

RESULT 8

LOCUS AQ160513/c

DEFINITION clone mgxb0006C21r, genomic survey sequence.

ACCESSION AQ160513

VERSION AQ160513.1 GI:3557502

KEYWORDS GSS.

SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)

ORGANISM Magnaporthe grisea

REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes Incertae sedis; Magnaportheaceae; Magnaporthe.

AUTHORS Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasnowski, M., Wing, R.A. and Dean, R.A.

TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome

JOURNAL Unpublished (1998)

COMMENT Contact: Dean RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson University, Clemson, SC 29634  
Tel: 864 656 5737  
Fax: 864 656 4293

Seq primer: GGAACAGCTATGACCATG

Class: BAC ends

High quality sequence stop: 424.

Location/Qualifiers

1. .603

/organism="Magnaporthe grisea"

/mol\_type="genomic DNA"

/strain="70-15"

/db\_xref="taxon:148305"

/clone="mgxb0006C21r"

/tissue\_type="Protoplasts"

/lab\_host="E. coli DH10B"

/clone\_lib="CUGI Rice Blast BAC Library"

/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with



TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea

Genome

LOCUS A0160254/c

DEFINITION mgxb0003L19r CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0003L19r, genomic survey sequence.

ACCESSION A0160254

VERSION A0160254.1 GI:3557243

KEYWORDS GSS.

SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)

ORGANISM Magnaporthe grisea

REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.

AUTHORS 1 (bases 1 to 750)  
Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.

TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea

Genome

UNPUBLISHED (1998)

CONTACT: Dean RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson University, Clemson, SC 29634

Tel: 864 656 5737

Fax: 864 656 4293

Email: rdean@clemson.edu

Seq primer: GGAACAGCTATGACCATG

Class: BAC ends

High quality sequence stop: 321.

Location/Qualifiers

1..720

/organism="Magnaporthe grisea"

/mol\_type="genomic DNA"

/strain="70-15"

/db\_xref="taxon:148305"

/clone="mgxb0004B19r"

/tissue\_type="Protoplasts"

/lab\_host="E. coli DH10B"

/clone\_lib="CUGI Rice Blast BAC Library"

/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25x genome coverage. High density colony filters are available upon request."

## ORIGIN

Query Match 5.4%; Score 82.2; DB 8; Length 720;

Best Local Similarity 54.1%; Pred. No. 2e-08;

Matches 236; Conservative 0; Mismatches 189; Indels 11; Gaps 3;

QY 848 AAAGGTAAAAAATTCATGAAACACAAACACCAACCAAGTTGGTAACATGTCCATAAAC 907

DB 445 AAAGGAAAAAAGAGACTTAACAACAACAAACAAAAA-----CAGCCGCGTCATCAAC 391

QY 908 TAGCGAGCCAACTTCCAAACCAATGTAATGCGTATTTATGCGTCTATGTTGGACTGTT 967

DB 390 TACTCGGGCAGCTACAGCCNCAGGCAACTCATACCTGGCGCTCTACGGCTGGACGCGC 331

QY 968 GACCCCTTTCGATATTTATTTGTCGACAGTTGGGCAACTGGCGTCCACCA---GGA 1024

DB 330 AACCCGCTGATCGAGTACTACGTGGTGGAGAGCTTTGGCAGCTACAAACCGTCTGCGGC 271

QY 1025 GCAACGCTTAAGGGGACCATCTGTTGATGGAGGAACATATGATATCTACGAGACTCTT 1084

DB 270 GCCACCAACCGCGGTCTTCACCTCGAGCGGAGCACTACGACATCTCTGTCAGCACC 211

QY 1085 AGAGTCAATCAACCTCATTAAGGGGATGCCACATTTAAACAATATTGGAGTTTCGA 1144

DB 210 CGGTACAAACAGCCCTCCATCGAGCGGCAACAAAGACCTTTCAGCAGTCTCTGTCGTCGC 151

QY 1145 AGATCGAAACGACAGTGGGACGATTTCTGTCAGCAACCACTTTAGAGCGTGGGAAAC 1204

DB 150 CGCAACAAGCGCGGCGGACCGCTCACCTTTGCCAACCAACGTCACACGCTTGGCGGAAC 91

QY 1205 TTAGGATGAATATGGGGAATATGTAAGTTCGCG---CTTACTGTAGAAGGCTATCAA 1261

DB 90 GCCGCTCAACCTTCGGCAACCACTGGAATACAGATCTCTGGCGCTGAGGGCTACCAAC 31

QY 1262 AGTAGCGGAAGTGCTA 1277

DB 30 AGCAGCGGCTCCGCCA 15

RESULT 11

LOCUS A0160254/c

DEFINITION mgxb0003L19r CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0003L19r, genomic survey sequence.

ACCESSION A0160254

VERSION A0160254.1 GI:3557243

KEYWORDS GSS.

SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)

ORGANISM Magnaporthe grisea

REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.

AUTHORS 1 (bases 1 to 750)  
Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.

TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea

Genome

UNPUBLISHED (1998)

CONTACT: Dean RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson University, Clemson, SC 29634

Tel: 864 656 5737

Fax: 864 656 4293

Email: rdean@clemson.edu

Seq primer: GGAACAGCTATGACCATG

Class: BAC ends

High quality sequence stop: 344.

Location/Qualifiers

1..750

/organism="Magnaporthe grisea"

/mol\_type="genomic DNA"

/strain="70-15"

/db\_xref="taxon:148305"

/clone="mgxb0003L19r"

/tissue\_type="Protoplasts"

/lab\_host="E. coli DH10B"

/clone\_lib="CUGI Rice Blast BAC Library"

/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25x genome coverage. High density colony filters are available upon request."

## FEATURES

source

## ORIGIN

Query Match 5.4%; Score 82.2; DB 8; Length 750;

Best Local Similarity 54.1%; Pred. No. 2e-08;

Matches 236; Conservative 0; Mismatches 189; Indels 11; Gaps 3;

QY 848 AAAGGTAAAAAATTCATGAAACACAAACACCAACCAAGTTGGTAACATGTCCATAAAC 907

DB 445 AAAGGAAAAAAGAGACTTAACAACAACAAACAAAAA-----CAGCCGCGTCATCAAC 391

QY 908 TAGCGAGCCAACTTCCAAACCAATGTAATGCGTATTTATGCGTCTATGTTGGACTGTT 967

DB 390 TACTCGGGCAGCTACAGCCNCAGGCAACTCATACCTGGCGCTCTACGGCTGGACGCGC 331

QY 968 GACCCCTTTCGATATTTATTTGTCGACAGTTGGGCAACTGGCGTCCACCA---GGA 1024

DB 330 AACCCGCTGATCGAGTACTACGTGGTGGAGAGCTTTGGCAGCTACAAACCGTCTGCGGC 271

QY 1025 GCAACGCTTAAGGGGACCATCTGTTGATGGAGGAACATATGATATCTACGAGACTCTT 1084

DB 270 GCCACCAACCGCGGTCTTCACCTCGAGCGGAGCACTACGACATCTCTGTCAGCACC 211

QY 1085 AGAGTCAATCAACCTCATTAAGGGGATGCCACATTTAAACAATATTGGAGTTTCGA 1144

DB 210 CGGTACAAACAGCCCTCCATCGAGCGGCAACAAAGACCTTTCAGCAGTCTCTGTCGTCGC 151

QY 1145 AGATCGAAACGACAGTGGGACGATTTCTGTCAGCAACCACTTTAGAGCGTGGGAAAC 1204

DB 150 CGCAACAAGCGCGGCGGACCGCTCACCTTTGCCAACCAACGTCACACGCTTGGCGGAAC 91

QY 1205 TTAGGATGAATATGGGGAATATGTAAGTTCGCG---CTTACTGTAGAAGGCTATCAA 1261

DB 90 GCCGCTCAACCTTCGGCAACCACTGGAATACAGATCTCTGGCGCTGAGGGCTACCAAC 31

QY 1262 AGTAGCGGAAGTGCTA 1277

DB 30 AGCAGCGGCTCCGCCA 15



blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

# ORIGIN

```

Query Match      5.4%; Score 81.4; DB 8; Length 583;
Best Local Similarity 52.8%; Pred. No. 2.9e-08;
Matches 200; Conservative 0; Mismatches 176; Indels 3; Gaps 1;

QY 854 AAAAAATTCATGAACACAAACACACCAACAGTTGGTAACATGTCCATAAAGTACGGA 913
    |||||
Db 445 AAAAAAAGAGAGAGACTAACATCAACAACAACAAACAGCGCGTCACTACTCG 386

QY 914 GCCAACTTCCAAACCAATGGTATATGCGTCTATGCTGTTGGAAGTGTGACCT 973
    |||||
Db 385 GCGAGCTACAGCCACAGGCAACTCATCTGGCGTCTACGGTGGAGCGCAACCG 326

QY 974 CTGTGCAATATATATGTGACAGTTGGGGCACTGGCGTCCACCA---GGAGCAAG 1030
    |||||
Db 325 CTGATCGAGTACTACGTGGTGGAGAGCTTTGGCAGCTACAACCCGTCGTGGGCGCCACC 266

QY 1031 CCTAAGGGACCATCTGTTGATGAGGACATATGATCTACGAGACTCTTAGAGTC 1090
    |||||
Db 265 AACCAGCGCTCTTCACTCGGACGCGAGCAGCTACGACATCTGTCAGCAGCCGCTAC 206

QY 1091 AATCAACCTTCCATTAAGGGATGCGCACATTTAAACAATATGAGAGTGTGGAATCG 1150
    |||||
Db 205 AACCAGCGCTTCCATCGAGCGCAGCAGCTTTCAGCAGTCTGCTGGTGGCGGCAAC 146

QY 1151 AAGCGCAGGAGTGGCAGAGTTTCTGTCAGCAACCTTTAGAGCGTGGGAAAAGTGGG 1210
    |||||
Db 145 AAGCGCGCAGCGCAGCGTCTTGGCAACAGCTCAACGCTGGCGGCAACGCGCGC 86

QY 1211 ATGAATATGGGAAAATGT 1229
    |||||
Db 85 CTCACCTCGGCAACAGT 67
  
```

```

RESULT 14
CNS00B01/c
LOCUS
DEFINITION
  Drosophila melanogaster genome survey sequence 17 end of BAC #
  BACR24D09 of RPCI-98 library from Drosophila melanogaster (fruit
  fly), genomic survey sequence.
ACCESSION
  AL057419.1 GI:4937885
VERSION
  GSS.
SOURCE
  Drosophila melanogaster (fruit fly)
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
  1 (bases 1 to 1101)
REFERENCE
  Genoscope.
  Direct Submission
  Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
  - web : www.genoscope.cns.fr)
COMMENT
  Determination of this BAC-end sequence was carried out as part of a
  collaboration with the Berkeley Drosophila Genome Project (BDGP).
  The BDGP is constructing a physical map of the Drosophila
  melanogaster genome using these BACs. For further information
  please see http://www.fruitfly.org The BDGP Drosophila
  melanogaster BAC library was prepared by Kazuo Osoegawa and
  Aaron Mammosier in Pieter de Jong's laboratory in the Department of
  Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
  NY. The library is named RPCI-98 and was constructed by partial
  EcoRI digestion of Drosophila DNA provided by the BDGP from the
  isogenic strain y2; cn bw sp, the same strain used for the BDGP's
  
```

P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

```

source
  1..1101
  Location/Qualifiers
    /organism="Drosophila melanogaster"
    /mol_type="genomic DNA"
    /db_xref="taxon:7227"
    /clone="BACR24D09"
    /clone_lib="RPCI-98"
    /note="end : 17"
  
```

## ORIGIN

```

Query Match      5.3%; Score 79.8; DB 9; Length 1101;
Best Local Similarity 38.7%; Pred. No. 7.6e-08;
Matches 239; Conservative 75; Mismatches 296; Indels 7; Gaps 2;

QY 72 TCTTTATTTAAAGCTTAACCACTTGCATCTTATCAAGAACATCTCTTTATAGGAAGT 131
    |||||
Db 1097 TTTTMMTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1038

QY 132 CCCATTTGCAAGACATTAATAAATCTTTTCCCTATTTATCTTATCGCTTGCCT 191
    |||||
Db 1037 ATWATTAATATWATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 978

QY 192 TTAATTTGTAACCTTTTATTTAGTTTACGTGATGTTCCCTCATTCATACCATTAAT 251
    |||||
Db 977 TTAATTTTAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 919

QY 252 GTTAAAGCTAGAGTCACTCTTTTTCGGTTCTCAAAATAACCTGAAAGAACATTAAT 311
    |||||
Db 918 ATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 859

QY 312 ATTTTCTCAGCGCGCTCCATTAATGAATATATATATCTCTTTTATACATTAATTAAT 371
    |||||
Db 858 WATTAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 799

QY 372 AGTATATACCTCGGTTATCAAAATGTAGATAATCAATTAATGATCAAAACAGCAGCT 431
    |||||
Db 798 TTTTMMTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 739

QY 432 AAAAAACACTGATGTGACCTCTTAAAGAAAGTGTCTATCTATGATGAAAGATAATTAAT 489
    |||||
Db 738 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 679

QY 490 ----CCAGTTTCAAAATTTGAAATAGTGTGATGATGATGATGATGATGATGATGAT 545
    |||||
Db 678 AATSVCASSSVMTTATATATTTTASAAAATAAATAAATAAATAAATAAATAAATAAATAA 619

QY 546 AAAGGAGGCTAGTAGTACGCTAGCTTCAATACCAAAATAATAGTTGTAATAAATAAATA 605
    |||||
Db 618 SSATAATTAATTSVTRATGTTDAAWTTTAAASAAAATAAATAAATAAATAAATAAATAAATA 559

QY 606 AGCAGGAATCCCTTAATGACAAAGAAATGACGTTGATTTTACGCTTTTATAGTTTGT 665
    |||||
Db 558 ARSATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 499

QY 666 TTGCACTAACCTTACCT 682
    |||||
Db 498 ATTTNNNNNNNTTGST 482
  
```

```

RESULT 15
AQ396475/c
LOCUS
DEFINITION
  clone mgxb0010M14f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
  clone mgxb0010M14f, genomic survey sequence.
ACCESSION
  AQ396475
VERSION
  AQ396475.1 GI:4367502
KEYWORDS
  Magnaporthe grisea (anamorph: Pyricularia grisea)
SOURCE
  Magnaporthe grisea
  ORGANISM
  
```

AQ396475 561 bp DNA linear GSS 06-MAR-1999  
 mgxb0010M14f CUGI Rice Blast BAC Library Magnaporthe grisea genomic  
 clone mgxb0010M14f, genomic survey sequence.



REFERENCE  
AUTHORS Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,  
Phillips, K., Sasnowski, M., Wing, R.A. and Dean, R.A.  
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea  
Genome  
JOURNAL Unpublished (1998)  
COMMENT Contact: Dean RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson University, Clemson, SC 29634  
Tel: 864 656 5737  
Fax: 864 656 4293  
Email: rdean@clemson.edu  
Seq primer: TAATAGCTACTATAGGG  
Class: BAC ends  
High quality sequence start: 50  
High quality sequence stop: 443.

FEATURES  
source

1..561  
/organism="Magnaporthe grisea"  
/mol\_type="genomic DNA"  
/strain="70-15"  
/db\_xref="taxon:148305"  
/clone="mgxb0010M14f"  
/tissue\_type="Protoplasts"  
/lab\_host="E. coli DH10B"  
/clone\_lib="CUGI Rice Blast BAC Library"  
/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;  
Rice blast is one of the most devastating fungal diseases  
of rice world wide. It is a filamentous ascomycete with  
a haploid genome (n=7) of approximately 40 Mbp. Rice  
blast is an important model fungal pathogen for studying  
numerous aspects of the fungal-host interaction. In  
order to facilitate genome wide analysis, a BAC library  
containing 9216 clones with an average insert size of 130  
kbp was constructed. This library represents greater  
than 25X genome coverage. High density colony filters  
are available upon request."

## ORIGIN

Query Match 5.1%; Score 76.6; DB 8; Length 561;  
Best Local Similarity 53.7%; Pred. No. 3.9e-07;  
Matches 204; Conservative 0; Mismatches 170; Indels 6; Gaps 2;  
QY 839 TTATTCCGTAAAGTAAATAATTCATGAACACAAACACACCAACAGTTGGTAACATG 898  
Db 495 TTTTCCCAAAAGGAAAGAAAGAGACTANCAACACCAACAAACAGCGCGGT- 437  
QY 899 TCATAAACTACGGAGCCAACTTCCAAACCAATGGTAATGCTATTTATGGTCTATGGT 958  
Db 436 --CATCACTACTCGGGCAGCTAGAGCCCGGCAACTATACCTGGCGGTCTACGGC 379  
QY 959 TGGACTGTGACCCCTTTGTGGAATATTATATTGTCGACAGTTGGGGCAACTGGCGTCCA 1018  
Db 378 TGGACCGCGCAACCCGCTGATCGAGTACTACGTGGTGGAGAGCTTTGGCAGCTACAAACCG 319  
QY 1019 CCA---GGAGCAACCCCTAAGGGGACCATCACTGTTGATGGAGGACATATGATATCTAC 1075  
Db 318 TCGTCGGGGCGCCACCACACCGCGGGTCTTTCACCTCGGACGGGACCATCGACATCTG 259  
QY 1076 GAGACTCTTAGAGTCAATCAACCCCTCCATTAAAGGGGATTCACATTTAAACAATATTGG 1135  
Db 258 GTCAGCACCCCGCTACAACAGCCCTCCATCGAGCGCACCAAGACCTTTCAGCAGTTCTGG 199  
QY 1136 AGTGTTCGAAGATCGAAGCGACAGTGGCGAGATTCTGTGAGCAACCACTTTAGAGCG 1195  
Db 198 TCGGTGCGCGCGCAACAAGCGCGCACGCGCACCGTCACTTTGGCCAAACCAAGTCAACGCC 139  
QY 1196 TGGGAAACTTTAGGGATGAA 1215  
Db 138 TGGCGCAACGCGGCTCAA 119



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 10, 2004, 15:57:32 ; Search time 894.922 Seconds  
(without alignments)  
9775.823 Million cell updates/sec

Title: US-09-909-207-26  
Perfect score: 185  
Sequence: 1 TCATGTAACCTGCCTTGATC.....AAAAGAGAGGATGGA 185

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hgt.\*  
3: gb\_in.\*  
4: gb\_cm.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	184	99.5	185	6 A42278	A42278 Sequence 28
2	184	99.5	185	6 A48247	A48247 Sequence 26
3	184	99.5	185	6 AR127043	AR127043 Sequence
4	184	99.5	185	6 AR193071	AR193071 Sequence
5	184	99.5	185	6 AR220047	AR220047 Sequence
6	184	99.5	185	6 AR221330	AR221330 Sequence 1
7	184	99.5	1022	6 A42251	A42251 Sequence 1
8	184	99.5	1022	6 A42285	A42285 Sequence 35
9	184	99.5	1022	6 AR127019	AR127019 Sequence
10	184	99.5	1022	6 AR127049	AR127049 Sequence
11	184	99.5	1022	6 AR220023	AR220023 Sequence
12	184	99.5	1022	6 AR220053	AR220053 Sequence
13	184	99.5	1022	6 AR221306	AR221306 Sequence
14	184	99.5	1022	6 AR221336	AR221336 Sequence
15	77.2	41.7	1789	1 AF326785	AF326785 Bacillus
16	39.8	21.5	176689	5 BX000477	BX000477 Zebrafish
17	38.6	20.9	158951	9 AL355390	AL355390 Human DNA
18	37.6	20.3	178560	10 AC134839	AC134839 Mus muscu
19	37.6	20.3	208557	10 AC125406	AC125406 Mus muscu

C	20	37.4	20.2	207337	2	BX324185	BX324185 Danio rer
C	21	37.2	20.1	105587	9	CNS01DU6	AL133166 Human chr
C	22	37	20.0	108566	2	AC098193_6	Continuation (7 of
C	23	37	20.0	110000	2	AC098192_1	Continuation (2 of
C	24	37	20.0	110000	2	AC098192_2	Continuation (3 of
C	25	37	20.0	120255	9	AL589167	AL589167 Human DNA
C	26	37	20.0	244759	2	AC098158	AC098158 Rattus no
C	27	37	20.0	253645	2	AC137382	AC137382 Rattus no
C	28	36.2	19.6	127249	2	CR383671	CR383671 Danio rer
C	29	36	19.5	215357	2	AC105716	AC105716 Rattus no
C	30	36	19.5	231967	2	AC111688	AC111688 Rattus no
C	31	35.8	19.4	7218	6	I66494	I66494 Sequence 14
C	32	35.6	19.2	233510	2	AC136569	AC136569 Rattus no
C	33	35.6	19.2	238007	10	AC097745	AC097745 Rattus no
C	34	35.6	19.2	246818	2	AC117880	AC117880 Rattus no
C	35	35.4	19.1	276	8	CNS01CFB	AL115455 Botrytis
C	36	35.4	19.1	420	8	CNS01DGI	AL116777 Botrytis
C	37	35.4	19.1	540	8	CNS01AIJ	AL112367 Botrytis
C	38	35.4	19.1	600	8	CNS01CYC	AL116140 Botrytis
C	39	35.4	19.1	97101	2	AL450308	AL450308 Homo sapi
C	40	35.4	19.1	130878	9	AL450063	AL450063 Human DNA
C	41	35.4	19.1	182862	2	AC137480	AC137480 Rattus no
C	42	35.4	19.1	200446	2	AC125593	AC125593 Rattus no
C	43	35.2	19.0	206981	2	BX316384	BX316384 Danio rer
C	44	35.2	19.0	287360	2	BX942821	BX942821 Danio rer
C	45	35	18.9	160439	9	AC005345	AC005345 Homo sapi

ALIGNMENTS

RESULT 1  
A42278  
LOCUS A42278  
DEFINITION Sequence 28 from Patent EP0634490.  
ACCESSION A42278  
VERSION A42278.1 GI:2297766  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 185)  
AUTHORS De,B.E., Lahaya,A., Ledoux,P., Amory,A., Detroz,R., Andre,C. and Vetter,R.  
TITLE Xylanase derived from a bacillus species, expression vectors for such xylanase and other proteins, host organisms therefor and use thereof  
JOURNAL Patent: EP 0634490-A 28 18-JAN-1995;  
COMMENT SOLVAY (BE)  
Other publication NZ 260989 950828  
Other publication BR 9402834 950613  
Other publication JP 7067637 950314  
Other publication FI 943389 950116  
Other publication CA 2128050 950116  
Other publication AU 942652 950116  
Other publication AU 6743294 950127  
Other publication GB 2279955 950118.

FEATURES

source 1..185  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

ORIGIN

Query Match 99.5%; Score 184; DB 6; Length 185;  
Best Local Similarity 100.0%; Pred. No. 2.2e-45;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TCATGTAACCTGCCTTGATCTATTTCATTTGTTATCAAGGATTTATACAAACAAGAGA 60  
Db 1 TCATGTAACCTGCCTTGATCTATTTCATTTGTTATCAAGGATTTATACAAACAAGAGA 60  
Qy 61 CATCCATGCCGGGTTAAAGCAGTATCGTTCATCTAAACAGAGAGGNCCTGCATGAAAGGA 120

Db 61 CATCCATCGCGGTTAAAGCAGTATCGTTCCATCTAACAGAGAAGGCTGCATGAAGGA 120  
 Qy 121 GGTGATGGGTTTTTCATCTTAGGATGACAGAACATACGGATGAAAAAAGGAGAGGAT 180  
 Db 121 GGTGATGGGTTTTTCATCTTAGGATGACAGAACATACGGATGAAAAAAGGAGAGGAT 180  
 Qy 181 GGAAA 185  
 Db 181 GGAAA 185

RESULT 2  
 A48247  
 LOCUS Sequence 26 from Patent EP0698667. 185 bp DNA linear PAT 07-MAR-1997  
 DEFINITION A48247  
 ACCESSION A48247  
 VERSION A48247.1 GI:2302095  
 KEYWORDS  
 SOURCE unidentified  
 ORGANISM unidentified  
 UNCLASSIFIED.  
 REFERENCE 1 (bases 1 to 185)  
 AUTHORS De, B. E., Lahaye, A., Ledoux, P. and Detroz, R.  
 TITLE Xylanase, microorganisms for its production, DNA molecules, process of preparation and use thereof  
 JOURNAL Patent: EP 0698667-A 26 28-FEB-1996;  
 SOLVAY (BE)  
 COMMENT Other publication BE 1008751 960702  
 Other publication BE 1008570 960604  
 Other publication BR 9503454 960305  
 Other publication JP 8092284 960409  
 Other publication FI 953578 960127  
 Other publication CA 2154628 960127  
 Other publication AU 2508695 960208.

FEATURES  
 source  
 1..185  
 /organism="unidentified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"

ORIGIN

Query Match 99.5%; Score 184; DB 6; Length 185;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-45;  
 Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCATGTAACCTCGCTTGATCTATTTCATTTGATCAAGGATTTATACAAACAAGAGA 60  
 Db 1 TCATGTAACCTCGCTTGATCTATTTCATTTGATCAAGGATTTATACAAACAAGAGA 60  
 Qy 61 CATCCATCGCGGTTAAAGCAGTATCGTTCCATCTAACAGAGAAGGCTGCATGAAGGA 120  
 Db 61 CATCCATCGCGGTTAAAGCAGTATCGTTCCATCTAACAGAGAAGGCTGCATGAAGGA 120  
 Qy 121 GGTGATGGGTTTTTCATCTTAGGATGACAGAACATACGGATGAAAAAAGGAGAGGAT 180  
 Db 121 GGTGATGGGTTTTTCATCTTAGGATGACAGAACATACGGATGAAAAAAGGAGAGGAT 180  
 Qy 181 GGAAA 185  
 Db 181 GGAAA 185

RESULT 3  
 AR127043  
 LOCUS Sequence 28 from patent US 6180382. 185 bp DNA linear PAT 16-MAY-2001  
 DEFINITION AR127043  
 ACCESSION AR127043  
 VERSION AR127043.1 GI:14113636  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 185)  
 AUTHORS De Buyl, E., Lahaye, A., Ledoux, P., Amory, A., Detroz, R., Andre, C. and Vetter, R.  
 TITLE Xylanase derived from a bacillus species, expression vectors for such xylanase and other proteins, host organisms therefor and use thereof  
 JOURNAL Patent: US 6180382-A 28 30-JAN-2001;  
 FEATURES Location/Qualifiers  
 source 1..185  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

Query Match 99.5%; Score 184; DB 6; Length 185;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-45;  
 Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCATGTAACCTCGCTTGATCTATTTCATTTGATCAAGGATTTATACAAACAAGAGA 60  
 Db 1 TCATGTAACCTCGCTTGATCTATTTCATTTGATCAAGGATTTATACAAACAAGAGA 60  
 Qy 61 CATCCATCGCGGTTAAAGCAGTATCGTTCCATCTAACAGAGAAGGCTGCATGAAGGA 120  
 Db 61 CATCCATCGCGGTTAAAGCAGTATCGTTCCATCTAACAGAGAAGGCTGCATGAAGGA 120  
 Qy 121 GGTGATGGGTTTTTCATCTTAGGATGACAGAACATACGGATGAAAAAAGGAGAGGAT 180  
 Db 121 GGTGATGGGTTTTTCATCTTAGGATGACAGAACATACGGATGAAAAAAGGAGAGGAT 180  
 Qy 181 GGAAA 185  
 Db 181 GGAAA 185

RESULT 4  
 AR193071  
 LOCUS Sequence 26 from patent US 6346407. 185 bp DNA linear PAT 20-APR-2002  
 DEFINITION AR193071  
 ACCESSION AR193071  
 VERSION AR193071.1 GI:20239036  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 185)  
 AUTHORS De Buyl, E., Lahaye, A., Ledoux, P. and Detroz, R.  
 TITLE Xylanase, microorganisms producing it, DNA molecules, methods for preparing this xylanase and uses of the latter  
 JOURNAL Patent: US 6346407-A 26 12-FEB-2002;  
 FEATURES Location/Qualifiers  
 source 1..185  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

Query Match 99.5%; Score 184; DB 6; Length 185;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-45;  
 Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCATGTAACCTCGCTTGATCTATTTCATTTGATCAAGGATTTATACAAACAAGAGA 60  
 Db 1 TCATGTAACCTCGCTTGATCTATTTCATTTGATCAAGGATTTATACAAACAAGAGA 60  
 Qy 61 CATCCATCGCGGTTAAAGCAGTATCGTTCCATCTAACAGAGAAGGCTGCATGAAGGA 120  
 Db 61 CATCCATCGCGGTTAAAGCAGTATCGTTCCATCTAACAGAGAAGGCTGCATGAAGGA 120  
 Qy 121 GGTGATGGGTTTTTCATCTTAGGATGACAGAACATACGGATGAAAAAAGGAGAGGAT 180  
 Db 121 GGTGATGGGTTTTTCATCTTAGGATGACAGAACATACGGATGAAAAAAGGAGAGGAT 180  
 Qy 181 GGAAA 185  
 Db 181 GGAAA 185

Db 181 GGAAA 185

RESULT 5  
AR220047 LOCUS 185 bp DNA linear PAT 26-SEP-2002  
DEFINITION Sequence 28 from patent US 6423523.  
ACCESSION AR220047  
VERSION AR220047.1 GI:23324465  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 185)  
AUTHORS de Buyt, E., Lahaye, A., Ledoux, P., Amory, A., Detroz, R., Andre, C. and Vetter, R.  
TITLE Xylanase derived from a bacillus species, expression vectors for such xylanase and other proteins, host organisms therefor and use thereof  
JOURNAL Patent: US 6423523-A 28 JUL-2002;  
FEATURES Location/Qualifiers  
1..185  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 99.5%; Score 184; DB 6; Length 185;  
Best Local Similarity 100.0%; Pred. No. 2.2e-45;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATGTAACCTCGCCTTGATCTATTTCATTTGTATCAAAAGGATTTATACACAAACAAGAGA 60  
DB 1 TCATGTAACCTCGCCTTGATCTATTTCATTTGTATCAAAAGGATTTATACACAAACAAGAGA 60

QY 61 CATCCATGCCGGTTAAAGCAGTAGTTCGTTCCATCTAACAGAGAGGNCCTGCATGAAGGA 120  
DB 61 CATCCATGCCGGTTAAAGCAGTAGTTCGTTCCATCTAACAGAGAGGNCCTGCATGAAGGA 120

QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACCGATGAAAAAGGAGAGGGAT 180  
DB 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACCGATGAAAAAGGAGAGGGAT 180

QY 181 GGAAA 185  
DB 181 GGAAA 185

RESULT 6  
AR221330 LOCUS 185 bp DNA linear PAT 26-SEP-2002  
DEFINITION Sequence 28 from patent US 6426211.  
ACCESSION AR221330  
VERSION AR221330.1 GI:23328305  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 185)  
AUTHORS de Buyt, E., Lahaye, A., Ledoux, P., Amory, A., Detroz, R., Andre, C. and Vetter, R.  
TITLE Xylanase derived from a Bacillus species, expression vectors for such xylanase and other proteins, host organisms therefor and use thereof  
JOURNAL Patent: US 6426211-A 28 JUL-2002;  
FEATURES Location/Qualifiers  
1..185  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 99.5%; Score 184; DB 6; Length 185;  
Best Local Similarity 100.0%; Pred. No. 2.2e-45;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATGTAACCTCGCCTTGATCTATTTCATTTGTATCAAAAGGATTTATACACAAACAAGAGA 60  
DB 1 TCATGTAACCTCGCCTTGATCTATTTCATTTGTATCAAAAGGATTTATACACAAACAAGAGA 60

QY 61 CATCCATGCCGGTTAAAGCAGTAGTTCGTTCCATCTAACAGAGAGGNCCTGCATGAAGGA 120  
DB 61 CATCCATGCCGGTTAAAGCAGTAGTTCGTTCCATCTAACAGAGAGGNCCTGCATGAAGGA 120

QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACCGATGAAAAAGGAGAGGGAT 180  
DB 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACCGATGAAAAAGGAGAGGGAT 180

QY 181 GGAAA 185  
DB 181 GGAAA 185

QY 1 TCATGTAACCTCGCCTTGATCTATTTCATTTGTATCAAAAGGATTTATACACAAACAAGAGA 60  
DB 1 TCATGTAACCTCGCCTTGATCTATTTCATTTGTATCAAAAGGATTTATACACAAACAAGAGA 60

QY 61 CATCCATGCCGGTTAAAGCAGTAGTTCGTTCCATCTAACAGAGAGGNCCTGCATGAAGGA 120  
DB 61 CATCCATGCCGGTTAAAGCAGTAGTTCGTTCCATCTAACAGAGAGGNCCTGCATGAAGGA 120

QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACCGATGAAAAAGGAGAGGGAT 180  
DB 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACCGATGAAAAAGGAGAGGGAT 180

QY 181 GGAAA 185  
DB 181 GGAAA 185

RESULT 7  
A42251 LOCUS 1022 bp DNA linear PAT 05-MAR-1997  
DEFINITION Sequence 1 from Patent EP0634490.  
ACCESSION A42251  
VERSION A42251.1 GI:2297742  
KEYWORDS  
SOURCE Bacillus pumilus  
ORGANISM Bacillus pumilus  
REFERENCE 1 (bases 1 to 1022)  
AUTHORS De, B.E., Lahaya, A., Ledoux, P., Amory, A., Detroz, R., Andre, C. and Vetter, R.  
TITLE Xylanase derived from a bacillus species, expression vectors for such xylanase and other proteins, host organisms therefor and use thereof  
JOURNAL Patent: EP 0634490-A 1 18-JAN-1995;  
COMMENT SOLVAY (BE)  
Other publication NZ 260989 950828  
Other publication BR 9402834 950613  
Other publication JP 7067637 950314  
Other publication FI 943389 950116  
Other publication CA 2138050 950116  
Other publication NO 942652 950116  
Other publication AU 6743294 950127  
Other publication GB 2279955 950118.

FEATURES  
source Location/Qualifiers  
1..1022  
/organism="Bacillus pumilus"  
/mol\_type="unassigned DNA"  
/strain="PRL B12"  
/db\_xref="taxon:1408"

ORIGIN

Query Match 99.5%; Score 184; DB 6; Length 1022;  
Best Local Similarity 100.0%; Pred. No. 1.9e-45;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATGTAACCTCGCCTTGATCTATTTCATTTGTATCAAAAGGATTTATACACAAACAAGAGA 60  
DB 1 TCATGTAACCTCGCCTTGATCTATTTCATTTGTATCAAAAGGATTTATACACAAACAAGAGA 60

QY 61 CATCCATGCCGGTTAAAGCAGTAGTTCGTTCCATCTAACAGAGAGGNCCTGCATGAAGGA 120  
DB 61 CATCCATGCCGGTTAAAGCAGTAGTTCGTTCCATCTAACAGAGAGGNCCTGCATGAAGGA 120

QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACCGATGAAAAAGGAGAGGGAT 180  
DB 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACCGATGAAAAAGGAGAGGGAT 180

QY 181 GGAAA 185  
DB 181 GGAAA 185

```
RESULT 8
A42285          A42285          1022 bp      DNA      linear      PAT 05-MAR-1997
LOCUS
DEFINITION      Sequence 35 from Patent EP0634490.
ACCESSION       A42285
VERSION         A42285.1 GI:2297772
SOURCE          unidentified
ORGANISM         unidentified
unclassified.
REFERENCE        1 (bases 1 to 1022)
AUTHORS         De,B., Lahaya,A., Ledoux,P., Amory,A., Detroz,R., Andre,C. and
                Vetter,R.
TITLE           xylanase derived from a bacillus species, expression vectors for
                such xylanase and other proteins, host organisms therefor and use
                thereof
JOURNAL
PATENT: EP 0634490-A 35 18-JAN-1995;
SOLVAY (BE)
COMMENT          Other publication NZ 260989 950828
                Other publication BR 9402834 950613
                Other publication JP 7057637 950314
                Other publication FI 94389 950116
                Other publication CA 2128050 950116
                Other publication NO 942652 950116
                Other publication AU 6743294 950127
                Other publication GB 2279955 950118.
FEATURES        source
                1..1022
                /organism="unidentified"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32644"
ORIGIN
Query Match      99.5%; Score 184; DB 6; Length 1022;
Best Local Similarity 100.0%; Pred. No. 1.9e-45;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAAAGGATTTATACAAAACAAGAGA 60
    |||||||
DB 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAAAGGATTTATACAAAACAAGAGA 60
    |||||||
QY 61 CATCCATCCCGGGTTAAAGCAGTATCGTTCATCTTAACAGAGAAGGCTGCATGAAGGA 120
    |||||||
DB 61 CATCCATCCCGGGTTAAAGCAGTATCGTTCATCTTAACAGAGAAGGCTGCATGAAGGA 120
    |||||||
QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAAACAATACGGATGAAAAAAGGAGGGAT 180
    |||||||
DB 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAAACAATACGGATGAAAAAAGGAGGGAT 180
    |||||||
QY 181 GGAAA 185
    |||||
DB 181 GGAAA 185
    |||||
QY 181 GGAAA 185
    |||||
DB 181 GGAAA 185
    |||||
RESULT 9
AR127019        AR127019        1022 bp      DNA      linear      PAT 16-MAY-2001
LOCUS
DEFINITION      Sequence 1 from patent US 6180382.
ACCESSION       AR127019
VERSION         AR127019.1 GI:14113612
KEYWORDS        Unknown.
SOURCE          Unknown.
ORGANISM         Unknown.
unclassified.
REFERENCE        1 (bases 1 to 1022)
AUTHORS         De Buyl,E., Lahaye,A., Ledoux,P., Amory,A., Detroz,R., Andre,C. and
                Vetter,R.
TITLE           xylanase derived from a bacillus species, expression vectors for
                such xylanase and other proteins, host organisms therefor and use
                thereof
JOURNAL
PATENT: US 6180382-A 1 30-JAN-2001;
LOCATION/Qualifiers
FEATURES        source
                1..1022
                /organism="unassigned DNA"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32644"
ORIGIN
Query Match      99.5%; Score 184; DB 6; Length 1022;
Best Local Similarity 100.0%; Pred. No. 1.9e-45;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAAAGGATTTATACAAAACAAGAGA 60
    |||||||
DB 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAAAGGATTTATACAAAACAAGAGA 60
    |||||||
QY 61 CATCCATCCCGGGTTAAAGCAGTATCGTTCATCTTAACAGAGAAGGCTGCATGAAGGA 120
    |||||||
DB 61 CATCCATCCCGGGTTAAAGCAGTATCGTTCATCTTAACAGAGAAGGCTGCATGAAGGA 120
    |||||||
QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAAACAATACGGATGAAAAAAGGAGGGAT 180
    |||||||
DB 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAAACAATACGGATGAAAAAAGGAGGGAT 180
    |||||||
QY 181 GGAAA 185
    |||||
DB 181 GGAAA 185
    |||||
QY 181 GGAAA 185
    |||||
DB 181 GGAAA 185
    |||||
RESULT 11
AR220023        AR220023        1022 bp      DNA      linear      PAT 26-SEP-2002
LOCUS
DEFINITION      Sequence 1 from patent US 6423523.
FEATURES        source
                1..1022
                /organism="unassigned DNA"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32644"
ORIGIN
Query Match      99.5%; Score 184; DB 6; Length 1022;
Best Local Similarity 100.0%; Pred. No. 1.9e-45;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAAAGGATTTATACAAAACAAGAGA 60
    |||||||
DB 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAAAGGATTTATACAAAACAAGAGA 60
    |||||||
QY 61 CATCCATCCCGGGTTAAAGCAGTATCGTTCATCTTAACAGAGAAGGCTGCATGAAGGA 120
    |||||||
DB 61 CATCCATCCCGGGTTAAAGCAGTATCGTTCATCTTAACAGAGAAGGCTGCATGAAGGA 120
    |||||||
QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAAACAATACGGATGAAAAAAGGAGGGAT 180
    |||||||
DB 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAAACAATACGGATGAAAAAAGGAGGGAT 180
    |||||||
QY 181 GGAAA 185
    |||||
DB 181 GGAAA 185
    |||||
QY 181 GGAAA 185
    |||||
DB 181 GGAAA 185
    |||||
```

```
ORIGIN
Query Match      99.5%; Score 184; DB 6; Length 1022;
Best Local Similarity 100.0%; Pred. No. 1.9e-45;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAAAGGATTTATACAAAACAAGAGA 60
    |||||||
DB 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAAAGGATTTATACAAAACAAGAGA 60
    |||||||
QY 61 CATCCATCCCGGGTTAAAGCAGTATCGTTCATCTTAACAGAGAAGGCTGCATGAAGGA 120
    |||||||
DB 61 CATCCATCCCGGGTTAAAGCAGTATCGTTCATCTTAACAGAGAAGGCTGCATGAAGGA 120
    |||||||
QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAAACAATACGGATGAAAAAAGGAGGGAT 180
    |||||||
DB 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAAACAATACGGATGAAAAAAGGAGGGAT 180
    |||||||
QY 181 GGAAA 185
    |||||
DB 181 GGAAA 185
    |||||
QY 181 GGAAA 185
    |||||
DB 181 GGAAA 185
    |||||
RESULT 10
AR127049        AR127049        1022 bp      DNA      linear      PAT 16-MAY-2001
LOCUS
DEFINITION      Sequence 35 from patent US 6180382.
ACCESSION       AR127049
VERSION         AR127049.1 GI:14113642
KEYWORDS        Unknown.
SOURCE          Unknown.
ORGANISM         Unknown.
unclassified.
REFERENCE        1 (bases 1 to 1022)
AUTHORS         De Buyl,E., Lahaye,A., Ledoux,P., Amory,A., Detroz,R., Andre,C. and
                Vetter,R.
TITLE           xylanase derived from a bacillus species, expression vectors for
                such xylanase and other proteins, host organisms therefor and use
                thereof
JOURNAL
PATENT: US 6180382-A 35 30-JAN-2001;
LOCATION/Qualifiers
FEATURES        source
                1..1022
                /organism="unassigned DNA"
                /mol_type="unassigned DNA"
ORIGIN
Query Match      99.5%; Score 184; DB 6; Length 1022;
Best Local Similarity 100.0%; Pred. No. 1.9e-45;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAAAGGATTTATACAAAACAAGAGA 60
    |||||||
DB 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAAAGGATTTATACAAAACAAGAGA 60
    |||||||
QY 61 CATCCATCCCGGGTTAAAGCAGTATCGTTCATCTTAACAGAGAAGGCTGCATGAAGGA 120
    |||||||
DB 61 CATCCATCCCGGGTTAAAGCAGTATCGTTCATCTTAACAGAGAAGGCTGCATGAAGGA 120
    |||||||
QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAAACAATACGGATGAAAAAAGGAGGGAT 180
    |||||||
DB 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAAACAATACGGATGAAAAAAGGAGGGAT 180
    |||||||
QY 181 GGAAA 185
    |||||
DB 181 GGAAA 185
    |||||
QY 181 GGAAA 185
    |||||
DB 181 GGAAA 185
    |||||
RESULT 11
AR220023        AR220023        1022 bp      DNA      linear      PAT 26-SEP-2002
LOCUS
DEFINITION      Sequence 1 from patent US 6423523.
FEATURES        source
                1..1022
                /organism="unassigned DNA"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32644"
ORIGIN
Query Match      99.5%; Score 184; DB 6; Length 1022;
Best Local Similarity 100.0%; Pred. No. 1.9e-45;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAAAGGATTTATACAAAACAAGAGA 60
    |||||||
DB 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAAAGGATTTATACAAAACAAGAGA 60
    |||||||
QY 61 CATCCATCCCGGGTTAAAGCAGTATCGTTCATCTTAACAGAGAAGGCTGCATGAAGGA 120
    |||||||
DB 61 CATCCATCCCGGGTTAAAGCAGTATCGTTCATCTTAACAGAGAAGGCTGCATGAAGGA 120
    |||||||
QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAAACAATACGGATGAAAAAAGGAGGGAT 180
    |||||||
DB 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAAACAATACGGATGAAAAAAGGAGGGAT 180
    |||||||
QY 181 GGAAA 185
    |||||
DB 181 GGAAA 185
    |||||
QY 181 GGAAA 185
    |||||
DB 181 GGAAA 185
    |||||
```

ACCESSION AR220023  
VERSION AR220023.1 GI:23324441  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1022)  
AUTHORS de Buyl, E., Lahaye, A., Ledoux, P., Amory, A., Detroz, R., Andre, C. and Vetter, R.  
TITLE Xylanase derived from a bacillus species, expression vectors for such xylanase and other proteins, host organisms therefor and use thereof  
JOURNAL Patent: US 6423523-A 1 23-JUL-2002;  
FEATURES  
source Location/Qualifiers  
1. .1022  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 99.5%; Score 184; DB 6; Length 1022;  
Best Local Similarity 100.0%; Pred. No. 1.9e-45;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAGGATTTATACAAACAAGAGA 60  
DB 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAGGATTTATACAAACAAGAGA 60  
QY 61 CATCCATGCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAAGNCTGCATGAAGGA 120  
DB 61 CATCCATGCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAAGNCTGCATGAAGGA 120  
QY 121 GGTGATGGGTTTTTCATCTTAGGATGACAGACAATACGATGAAAGGAGAGGGAT 180  
DB 121 GGTGATGGGTTTTTCATCTTAGGATGACAGACAATACGATGAAAGGAGAGGGAT 180  
QY 181 GGAAA 185  
DB 181 GGAAA 185  
RESULT 12  
AR220053  
LOCUS AR220053 35 from patent US 6423523.  
DEFINITION  
ACCESSION AR220053  
VERSION AR220053.1 GI:23324471  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1022)  
AUTHORS de Buyl, E., Lahaye, A., Ledoux, P., Amory, A., Detroz, R., Andre, C. and Vetter, R.  
TITLE Xylanase derived from a bacillus species, expression vectors for such xylanase and other proteins, host organisms therefor and use thereof  
JOURNAL Patent: US 6423523-A 35 23-JUL-2002;  
FEATURES  
source Location/Qualifiers  
1. .1022  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 99.5%; Score 184; DB 6; Length 1022;  
Best Local Similarity 100.0%; Pred. No. 1.9e-45;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAGGATTTATACAAACAAGAGA 60  
DB 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAGGATTTATACAAACAAGAGA 60  
QY 61 CATCCATGCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAAGNCTGCATGAAGGA 120  
DB 61 CATCCATGCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAAGNCTGCATGAAGGA 120  
QY 121 GGTGATGGGTTTTTCATCTTAGGATGACAGACAATACGATGAAAGGAGAGGGAT 180  
DB 121 GGTGATGGGTTTTTCATCTTAGGATGACAGACAATACGATGAAAGGAGAGGGAT 180  
QY 181 GGAAA 185  
DB 181 GGAAA 185  
RESULT 14  
AR221336  
LOCUS AR221336 35 from patent US 6426211.  
DEFINITION  
ACCESSION AR221336  
VERSION AR221336.1 GI:23328311  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1022)  
AUTHORS de Buyl, E., Lahaye, A., Ledoux, P., Amory, A., Detroz, R., Andre, C. and Vetter, R.  
TITLE Xylanase derived from a bacillus species, expression vectors for such xylanase and other proteins, host organisms therefor and use thereof  
JOURNAL Patent: US 6426211-A 35 30-JUL-2002;  
FEATURES  
source Location/Qualifiers  
1. .1022  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 99.5%; Score 184; DB 6; Length 1022;  
Best Local Similarity 100.0%; Pred. No. 1.9e-45;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAGGATTTATACAAACAAGAGA 60  
DB 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAGGATTTATACAAACAAGAGA 60  
QY 61 CATCCATGCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAAGNCTGCATGAAGGA 120  
DB 61 CATCCATGCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAAGNCTGCATGAAGGA 120  
QY 121 GGTGATGGGTTTTTCATCTTAGGATGACAGACAATACGATGAAAGGAGAGGGAT 180  
DB 121 GGTGATGGGTTTTTCATCTTAGGATGACAGACAATACGATGAAAGGAGAGGGAT 180  
QY 181 GGAAA 185  
DB 181 GGAAA 185

Db 61 CATCCATGCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAAGNCTGCATGAAGGA 120  
QY 121 GGTGATGGGTTTTTCATCTTAGGATGACAGACAATACGATGAAAGGAGAGGGAT 180  
DB 121 GGTGATGGGTTTTTCATCTTAGGATGACAGACAATACGATGAAAGGAGAGGGAT 180  
QY 181 GGAAA 185  
DB 181 GGAAA 185  
RESULT 13  
AR221306  
LOCUS AR221306 1022 bp DNA linear PAT 26-SEP-2002  
DEFINITION  
ACCESSION AR221306  
VERSION AR221306.1 GI:23328281  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1022)  
AUTHORS de Buyl, E., Lahaye, A., Ledoux, P., Amory, A., Detroz, R., Andre, C. and Vetter, R.  
TITLE Xylanase derived from a Bacillus species, expression vectors for such xylanase and other proteins, host organisms therefor and use thereof  
JOURNAL Patent: US 6426211-A 1 30-JUL-2002;  
FEATURES  
source Location/Qualifiers  
1. .1022  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 99.5%; Score 184; DB 6; Length 1022;  
Best Local Similarity 100.0%; Pred. No. 1.9e-45;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAGGATTTATACAAACAAGAGA 60  
DB 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAGGATTTATACAAACAAGAGA 60  
QY 61 CATCCATGCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAAGNCTGCATGAAGGA 120  
DB 61 CATCCATGCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAAGNCTGCATGAAGGA 120  
QY 121 GGTGATGGGTTTTTCATCTTAGGATGACAGACAATACGATGAAAGGAGAGGGAT 180  
DB 121 GGTGATGGGTTTTTCATCTTAGGATGACAGACAATACGATGAAAGGAGAGGGAT 180  
QY 181 GGAAA 185  
DB 181 GGAAA 185  
RESULT 14  
AR221336  
LOCUS AR221336 1022 bp DNA linear PAT 26-SEP-2002  
DEFINITION  
ACCESSION AR221336  
VERSION AR221336.1 GI:23328311  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1022)  
AUTHORS de Buyl, E., Lahaye, A., Ledoux, P., Amory, A., Detroz, R., Andre, C. and Vetter, R.  
TITLE Xylanase derived from a Bacillus species, expression vectors for such xylanase and other proteins, host organisms therefor and use thereof  
JOURNAL Patent: US 6426211-A 35 30-JUL-2002;  
FEATURES  
source Location/Qualifiers  
1. .1022  
/organism="unknown"  
/mol\_type="genomic DNA"

```

source
1. .1022
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match          99.5%; Score 184; DB 6; Length 1022;
Best Local Similarity 100.0%; Pred. No. 1.9e-45;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATGTAACCTCGCTTGATCTATTTCATTTGTATCAAGGATTTATACACAAACAAGAGA 60
Db 1 TCATGTAACCTCGCTTGATCTATTTCATTTGTATCAAGGATTTATACACAAACAAGAGA 60

QY 61 CATCCATGCCGGGTTAAAGCAGTATCGTTCCATCTAACAGAGAGGCTGCGATGAAGGA 120
Db 61 CATCCATGCCGGGTTAAAGCAGTATCGTTCCATCTAACAGAGAGGCTGCGATGAAGGA 120

QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACCAATACGGATGAAAAAGGAGGGAT 180
Db 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACCAATACGGATGAAAAAGGAGGGAT 180

QY 181 GGAAA 185
Db 181 GGAAA 185

```

```

RESULT 15
AF326785
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .1789
/organism="Bacillus pumilus"
/mol_type="genomic DNA"
/db_xref="taxon:1408"
563. .1249
/codon_start=1
/transl_table=11
/product="endo-1,4-xylanase"
/protein_id="AAG61087.1"
/db_xref="GI:12620815"
/translation="WNLRKRLLLFVWCIGLTLILPAVPAHARTITNEMGNQSGYDYE
LWKDYGNTSMTLNNGGAFSSGNIGNALFRKKKFDSTRTTHQLGNI SINYNASFPN
GGSYLGVVWTFQSLAIFYIVDSWGTVPPTCAVKGSPYADGTYDIYETTLVNQPSI
IGIATPKOYLSVRQTKRTSGTIVSVAHFPRKWSLCKPMGKMYETAFTVEGYQSSGAN
VMTNQLFIGN"

CDS

```

```

ORIGIN
Query Match          41.7%; Score 77.2; DB 1; Length 1789;
Best Local Similarity 77.2%; Pred. No. 9.2e-13;
Matches 132; Conservative 0; Mismatches 34; Indels 5; Gaps 3;

QY 16 TCATCTATTTCATTTGTTATCAAGGATTTATACACAAACAAGAGACATCCATCGCGGTT 75
Db 394 TCATCTCTTTCTTTTGTATCAATCATTTATACACACAAAGAGACATTCATGCGGGGCT 453

```

```

QY 76 AAAGCAGTATCGTTCCATCTTAACAGAGAGGNCCTGCATGAAAGGAGGTGATGGTT--TT 133
Db 454 AGCAGATATCGAT--AATCTAATAGTAAAGTGGGCGCATGAAGGAGGTGATGAATTCGTT 511

QY 134 TCATCTTAGGATGACAGAACCAATACCGATGAAAAAGGAGGAGGATGAA 184
Db 512 TCATCTTAGAGATGACAGAAATCAAAAGGATG-AAAAAGGAGAGGAATGAA 561

```

Search completed: November 10, 2004, 21:26:34  
Job time : 897.922 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 15:55:17 : Search time 117.352 Seconds  
(without alignments)  
8275.436 Million cell updates/sec

Title: US-09-909-207-26  
Perfect score: 185  
Sequence: 1 TCATGTAACCTGCCTTGATC.....AAAAGGAGAGGATGGAAA 185

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 23Sep04: \*  
1: Geneseqn1980a: \*  
2: Geneseqn1990s: \*  
3: Geneseqn2000s: \*  
4: Geneseqn2001as: \*  
5: Geneseqn2001bs: \*  
6: Geneseqn2002as: \*  
7: Geneseqn2002bs: \*  
8: Geneseqn2003as: \*  
9: Geneseqn2003bs: \*  
10: Geneseqn2003cs: \*  
11: Geneseqn2003ds: \*  
12: Geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	184	99.5	1022	2	AaQ80923 B. pumilus
2	183	98.9	184	2	AaQ73996 Xylanase
3	32.2	17.4	23439	2	AaV74349 Staphyloc
4	32	17.3	405	8	Abx38967 Bovine ES
5	31.6	17.1	302	3	AaA1613 Human sec
6	31.6	17.1	28149	5	AbA17961 Human ner
7	31.6	17.1	32192	5	AbA17963 Human ner
8	31.6	17.1	110000	12	Continuation (5 of
9	31.6	17.1	110000	12	Continuation (6 of
10	31.4	17.0	110000	2	Continuation (5 of
11	31.2	16.9	443	4	AaI91410 Human pol
12	31	16.8	390	6	AbK39185 CDNA enco
13	31	16.8	390	8	AcA11514 Human lun
14	31	16.8	390	8	AcA02700 Lung can
15	31	16.8	390	10	AdH46742 Human lun
16	31	16.8	410	5	AbV39054 Human pro
17	31	16.8	2262	10	AdF00565 Bacterial
18	31	16.8	5124	4	AaK75882 Human imm
19	31	16.8	32183	4	AaS35994 Human car
20	31	16.8	32183	10	AdE46688 Human car
21	31	16.8	51935	4	AaK75883 Human imm

c	22	30.8	16.6	1132	6	ABZ15647
	23	30.8	16.6	262090	12	ADQ59207
	24	30.6	16.5	1238	5	AAS68017
	25	30.2	16.3	930	8	ACA28595
	26	30	16.2	831	10	ACA55420
	27	30	16.2	831	12	ADI55216
	28	30	16.2	15962	6	ABA01445
c	29	30	16.2	43795	3	AZ292583
c	30	29.8	16.1	6718	4	ABL09262
c	31	29.8	16.1	9147	4	ABL09206
c	32	29.8	16.1	178896	6	ABQ88146
	33	29.6	16.0	553	3	ACA44275
	34	29.6	16.0	610	9	ACH04273
	35	29.6	16.0	977	4	AAK91468
	36	29.6	16.0	977	5	AAAS40051
	37	29.6	16.0	977	9	ADB33011
c	38	29.6	16.0	1143	5	AAH65014
c	39	29.6	16.0	1323	8	ACA01823
c	40	29.6	16.0	1761	6	AAI47164
	41	29.6	16.0	4436	4	ABL07432
c	42	29.6	16.0	5669	6	ABA01094
c	43	29.6	16.0	8924	4	ABL07438
	44	29.6	16.0	110000	6	ABX08336_11
	45	29.6	16.0	110000	12	ADJ25985_11

## ALIGNMENTS

RESULT 1  
AAQ80923

ID AAQ80923 standard; DNA; 1022 BP.

XX AAQ80923;

DT 02-AUG-1995 (first entry)

XX B. pumilus xylanase gene.

XX Xylanase; endo-1,4-beta-D-xylanase; wood pulp; pulping; biobleaching;  
bleaching; Bacillus licheniformis; ss.

OS Bacillus pumilus.

EH Key Location/Qualifiers

FT 5'UTR 1..185

FT /tag= a

FT /note= "claim 12"

FT misc\_difference 107

FT /\*tag= b

FT /note= "base n at position 107 is not identified in the

FT specification"

FT 186..869

FT /\*tag= c

FT sig\_peptide

FT 186..266

FT /\*tag= d

FT /note= "claim 11"

FT mat\_peptide

FT 267..866

FT /\*tag= e

FT /EC\_number= "3.2.1.8"

FT /note= "claim 10"

XX GB2279955-A.

XX 18-JAN-1995.

PD 15-JUL-1993; 93GB-00014780.

XX (SOLV ) SOLVAY & CIE.

XX Amory A, Andre C, De Buyl E, Detroz R, Lahaye A, Ledoux P;

PI

XX WPI: 1995-039214/06.  
 DR P-PSDB; AAR68849.  
 XX  
 PT Purified xylanase from *Bacillus pumilus* PRL B12 - esp. produced in  
 PT transformed *Bacillus licheniformis*, and related DNA, vectors, etc., used  
 PT for pre-treatment of wood pulp to reduce chlorine or ozone consumption in  
 PT subsequent bleaching.  
 XX Claim 13; Fig 1a-1b; 97pp; English.  
 XX  
 CC A *Bacillus pumilus* PRL B12 (ATCC 55443) gene library was screened for  
 CC recombinant plasmids carrying the xylanase gene. A chromosomal fragment  
 CC obtained from isolate pBPX1 was subcloned and expressed in *Escherichia*  
 CC coli JM109. The sequence of the xylanase gene (as a 1022 bp Sau3AI  
 CC fragment) carried by a selected transformant is given in AAQ80923  
 XX  
 SQ Sequence 1022 BP; 329 A; 197 C; 229 G; 266 T; 0 U; 1 Other;  
 Query Match 99.5%; Score 184; DB 2; Length 1022;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-49; Mismatches 0; Indels 0; Gaps 0;  
 Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TCATGTAACCTCGCTTGATCTATTTCATTGTATCAAAAGGATTATACACAAACAAGAGA 60  
 Db 1 TCATGTAACCTCGCTTGATCTATTTCATTGTATCAAAAGGATTATACACAAACAAGAGA 60  
 QY 61 CATCATGCCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAAGGNCCTGCATGAAAGGA 120  
 Db 61 CATCATGCCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAAGGNCCTGCATGAAAGGA 120  
 QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACCGATGAAAAAGGAGGGAT 180  
 Db 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACCGATGAAAAAGGAGGGAT 180  
 QY 181 GGAAA 185  
 Db 181 GGAAA 185  
 RESULT 2  
 ID AAQ73996 standard; DNA; 184 BP.  
 XX  
 AC AAQ73996;  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 15-MAY-1996 (first entry)  
 XX  
 DE Xylanase promoter.  
 XX  
 KW Xylanase; thermostable enzyme; paper; pulp; bleaching; feedstuff; baking;  
 KW promoter; ss.  
 XX  
 OS *Bacillus pumilus*; strain PRL B12 (ATCC 55443).  
 XX  
 PN AU9525086-A.  
 XX  
 PD 08-FEB-1996.  
 XX  
 PF 19-JUL-1995; 95AU-00025086.  
 XX  
 PR 26-JUL-1994; 94BE-00000706.  
 PR 17-MAY-1995; 95BE-00000448.  
 XX  
 PA (SOLV ) SOLVAY SA.  
 XX  
 PI De Buyl E, Lahaye A, Ledoux P, Detroz R;  
 XX  
 DR WPI: 1996-117341/13.  
 XX  
 PT *Bacillus* derived xylanase active over wide pH range - used in treatment  
 PT of paper pulp, animal feeds and in bakery goods.

XX Claim 26; Page 69; 94pp; English.  
 XX  
 CC The *Bacillus pumilus* PRL B12 xylanase gene promoter (AAQ73996) and  
 CC presequence (AAQ73997) were incorporated into vectors used for the  
 CC expression of DNA (see AAT16101) coding for a mature, thermostable  
 CC xylanase (AAR92053) of *Bacillus* sp. 720/1. The enzyme can be produced in  
 CC transformed hosts, pref. *Bacillus licheniformis* or *B. pumilus*, for use in  
 CC the paper-pulp, animal feed and baking industries. (Updated on 16-OCT-  
 XX 2003 to standardise OS field)  
 SQ Sequence 184 BP; 63 A; 29 C; 45 G; 46 T; 0 U; 1 Other;  
 Query Match 98.9%; Score 183; DB 2; Length 184;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-49; Mismatches 0; Indels 0; Gaps 0;  
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TCATGTAACCTCGCTTGATCTATTTCATTGTATCAAAAGGATTATACACAAACAAGAGA 60  
 Db 1 TCATGTAACCTCGCTTGATCTATTTCATTGTATCAAAAGGATTATACACAAACAAGAGA 60  
 QY 61 CATCATGCCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAAGGNCCTGCATGAAAGGA 120  
 Db 61 CATCATGCCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAAGGNCCTGCATGAAAGGA 120  
 QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACCGATGAAAAAGGAGGGAT 180  
 Db 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACCGATGAAAAAGGAGGGAT 180  
 QY 181 GGAA 184  
 Db 181 GGAA 184  
 RESULT 3  
 ID AAV74349 standard; DNA; 23439 BP.  
 XX  
 AC AAV74349;  
 XX  
 DT 16-MAR-1999 (first entry)  
 XX  
 DE *Staphylococcus aureus* contig SEQ ID #38.  
 XX  
 KW Computer readable medium; vaccine; *S. aureus* infection; immunodetection;  
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;  
 KW skin infection; surgical wound infection; scalded skin syndrome;  
 KW toxic shock syndrome; ds.  
 XX  
 OS *Staphylococcus aureus*.  
 XX  
 FH Key Location/Qualifiers  
 FT misc\_feature 481..540  
 FT /\*tag= a  
 FT /note= "these bases represent a line of missing text in  
 FT the sequence listing in the specification. They are  
 FT included to maintain the nucleotide numbering given in  
 FT the specification for this DNA sequence"  
 FT 2281..2340  
 FT misc\_feature /\*tag= b  
 FT /note= "these bases represent a line of missing text in  
 FT the sequence listing in the specification. They are  
 FT included to maintain the nucleotide numbering given in  
 FT the specification for this DNA sequence"  
 FT 4081..4140  
 FT misc\_feature /\*tag= c  
 FT /note= "these bases represent a line of missing text in  
 FT the sequence listing in the specification. They are  
 FT included to maintain the nucleotide numbering given in  
 FT the specification for this DNA sequence"  
 FT 5881..5940  
 FT misc\_feature /\*tag= d  
 FT /note= "these bases represent a line of missing text in



the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

7681..7740

/\*tag= e

/note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

9481..9540

/\*tag= f

/note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

11281..11340

/\*tag= g

/note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

13081..13140

/\*tag= h

/note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

14881..14940

/\*tag= i

/note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

16681..16740

/\*tag= j

/note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

18481..18540

/\*tag= k

/note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

20281..20340

/\*tag= l

/note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

22081..22140

/\*tag= m

/note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

EP786519-A2.

30-JUL-1997.

07-JAN-1997; 97EP-00100117.

05-JAN-1996; 96US-0009861P.

(HUMA-) HUMAN GENOME SCI INC.

Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA; WPI; 1997-374922/35.

Polynucleotide(s) and proteins derived from Staphylococcus aureus -

stored on computer readable medium and used in the production of anti-S. aureus vaccines.

Claim 1; Page 339-352; 3271pp; English.

This sequence represents one of 5191 Staphylococcus aureus DNA sequences of the invention. The DNA sequences are recorded on a computer readable medium, preferably selected from a floppy or hard disk, random access memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using the S. aureus DNA sequences allows putative functions to be assigned so that protein-encoding or regulatory regions of commercial, therapeutic or industrial importance can be obtained. Specifically, sequences which are likely to encode antigens have been identified and these polypeptides can be used in a vaccine composition against S. aureus infection. The polypeptides can also be used in a kit for the immunodetection of S. aureus in a sample. S. aureus is implicated in numerous human diseases, including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scalded skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments) are useful as primers or probes for isolating homologues of any of the S. aureus DNA sequences contained on the computer readable medium

Sequence 23439 BP; 7768 A; 3506 C; 4455 G; 6921 T; 0 U; 789 Other;

Query Match 17.4%; Score 32.2; DB 2; Length 23439;

Best Local Similarity 52.2%; Pred. No. 14;

Matches 70; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 44 TATACACAAACAGACATCCATCCGGGTTTAAAGCAGATATCGTTCCATCTAACACAGAGA 103

DB 11481 TATGTATATAACCGTGTGATGGACTTAGGCGTTCATTGTTTCAGTCGTTGGTGGGATAAC 11540

QY 104 AGNCTGCATGAAGAGGATGATGGGTTTTTCATCTTAGGATGACAGACATACCGAT 163

DB 11541 ACGACCAAAAGAAATTTACGAAACGTTTGTTCATTAATATGSAAGATTAAATGATAACGAT 11600

QY 164 GAAAAAGGAGAGG 177

DB 11601 AAAAAACGAGATG 11614

RESULT 4

ABX38967/c

ID ABX38967 standard; cDNA; 405 BP.

XX ABX38967;

AC AC

XX AC

DT 20-FEB-2003 (first entry)

XX 20-FEB-2003 (first entry)

DE Bovine EST associated with lactation/muscle/fat deposition #4132.

XX

KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;

KW muscle deposition; fat deposition; genome mapping; gene identification;

KW gene analysis; cattle breeding.

XX

OS Bos Taurus.

XX

XX US2002137139-A1.

XX

PD 26-SEP-2002.

XX

PF 24-SEP-2001; 2001US-00960352.

XX

PR 12-JAN-1999; 99US-0115707P.

PR 11-JAN-2000; 2000US-00480902.

XX

PA (BYAT/) BYATT J C.

PA (MATH/) MATHIALAGAN N.

PA (TAON/) TAO N.

PA (WARR/) WARREN W C.

XX

PI Byatt JC, Mathialagan N, Tao N, Warren WC;  
 XX WPI; 2003-110599/10.  
 XX New nucleic acid associated with lactation, and muscle and fat  
 PT deposition, useful for genome mapping, gene identification and analysis,  
 PT cattle breeding, or for genetically improving cattle.  
 XX  
 PS Claim 2: SEQ ID NO 4132; 245pp; English.  
 XX  
 CC The invention relates to a purified nucleic acid molecule associated with  
 CC lactation or muscle and fat deposition (designated LMFD), derived from  
 CC cattle, and the LMFD nucleic acid can specifically hybridise to a second  
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences  
 CC appearing as ABE34836-ABX49947, or complements of them. Also included are  
 CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic  
 CC acid linked to a promoter and a 3' non-translated sequence that  
 CC functions in the cell to cause termination of transcription and addition  
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and  
 CC (2) determining a level or pattern of a molecule in a bovine cell or  
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any  
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a  
 CC complementary nucleic acid molecule obtained from the bovine cell or  
 CC tissue, where hybridisation between the marker nucleic acid and the  
 CC complementary nucleic acid permits the detection of the molecule; and (b)  
 CC detecting the level or pattern of the complementary nucleic acid, where  
 CC the detection of the complementary nucleic acid is predictive of the  
 CC level or pattern of the molecule. The LMFD nucleic acid is used for  
 CC determining a level or pattern of a molecule in a bovine cell or tissue.  
 CC It is useful for genome mapping, gene identification and analysis, cattle  
 CC breeding, preparation of constructs for use in cattle gene expression, or  
 CC for genetically improving cattle. The present sequence is one of the  
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The  
 CC present sequence was not shown in the specification but was obtained in  
 CC electronic format from the USPTO web site:  
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139  
 XX  
 SQ Sequence 405 BP; 138 A; 63 C; 64 G; 139 T; 0 U; 1 Other;  
 Query Match 17.3%; Score 32; DB 8; Length 405;  
 Best Local Similarity 53.7%; Pred. No. 4.1;  
 Matches 65; Conservative 0; Mismatches 56; Indels 0; Gaps 0;  
 QY 21 TATTTCATTGTPATCAAGGATTTATACAAACAGAGACATCCATCGGGTTAAAGC 80  
 Db 122 TATTCTTTTAAATTAAGTTTAAAGTTTACTACAGAAATCGCTTCAACATGGA 63  
 QY 81 AGTATCGTTCCATCTAACAGAGAGGNCCTGCATGAAAGAGGTGATGGTTTTCATCTT 140  
 Db 62 AGCAATTTCTCTTCAAGATTTATACACCTGCATGGAAGAGGTGATTCCTTTACATTTA 3  
 QY 141 A 141  
 Db 2 A 2  
 RESULT 5  
 AAA41613/c  
 ID AAA41613 standard; cDNA; 302 BP.  
 XX  
 AC AAA41613;  
 XX  
 DT 21-AUG-2000 (first entry)  
 XX  
 DE Human secreted expressed sequence tag SEQ ID NO:353.  
 XX  
 KW Human; mouse; xenopus; rat; secreted expressed sequence tag; sEST;  
 KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
 KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;  
 KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;  
 KW antiviral; antidiabetic; antiaesthatic; vulnerary; antiparkinsonian;  
 KW antitumor; osteoprotective; neuroprotective; nootropic; antipsoriatic;  
 KW cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine;

KW autoimmune disorder; multiple sclerosis; allergic condition;  
 KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
 KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
 KW central nervous system disorder; Alzheimer's disease; stroke;  
 KW Parkinson's disease; Huntington's disease; coagulation disorder;  
 KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour;  
 KW infection; depression; psoriasis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200021990-A1.  
 XX  
 PD 20-APR-2000.  
 XX  
 PF 15-OCT-1999; 99WO-US024205.  
 XX  
 PR 15-OCT-1998; 98US-0104435P.  
 XX  
 PA (GEMY) GENETICS INST INC.  
 XX  
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M;  
 XX WPI; 2000-317937/27.  
 XX  
 DR Isolated polynucleotides, and encoded proteins, comprising secreted  
 PT expressed sequence tags (sESTs), useful for treating various disorders  
 PT such as autoimmune, infectious, and central nervous system disorders.  
 XX  
 CC Claim 1; Page 249; 618pp; English.  
 CC  
 CC AAA41261 to AAA43419 represent specifically claimed secreted expressed  
 CC sequence tags (sESTs), isolated from human, mouse, xenopus and rat tissue  
 CC sources. The sESTs can have a range of activities depending on the  
 CC tissues they were isolated from. The activities include: chemotactic;  
 CC proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic;  
 CC haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial;  
 CC antifungal; antiviral; antidiabetic; antiaesthatic; vulnerary; antitumor;  
 CC osteoprotective; neuroprotective; nootropic; antiparkinsonian; antipsoriatic;  
 CC cerebroprotective; anticonvulsant; and antidepressant. The sESTs can be  
 CC used for gene therapy and in vaccines. The sESTs are useful as probes for  
 CC the identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the sESTs. Proteins encoded by the sESTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumours, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA43420 to AAA43425 represent linker variants which are given  
 CC in the exemplification of the present invention  
 XX  
 SQ Sequence 302 BP; 67 A; 84 C; 62 G; 89 T; 0 U; 0 Other;

Query Match 17.1%; Score 31.6; DB 3; Length 302;  
 Best Local Similarity 53.8%; Pred. No. 5;  
 Matches 64; Conservative 0; Mismatches 55; Indels 0; Gaps 0;  
 QY 56 AGAGACATCCATCGCGGGTTAAAGCAGTATGTTCCATCTTAACAGAGAGGNCCTGATGA 115  
 Db 180 AGAGTTTAAATGGGCTCAAAAAAGTGGCTGCTGCCAACAGAGAAAAGCTCTAAA 121  
 QY 116 AAGGAGGTGATGGGTTTTCATCTTAGGGATGACAGAACATACGGATGAAAAAGGAG 174  
 Db 120 TAGAACCTGGCAGGAAATACATTTAAGAGTTGGGAGAGAGTAACCCAGATAAAGAGGAG 62

RESULT 6  
 ABA17961/c  
 ID ABA17961 standard; DNA; 28149 BP.  
 XX

AC	ABAI7961;	PR	12-SEP-2000;	2000US-0231968P.
XX		PR	14-SEP-2000;	2000US-0232397P.
DT	23-JAN-2002 (first entry)	PR	14-SEP-2000;	2000US-0232398P.
XX		PR	14-SEP-2000;	2000US-0232399P.
DE	Human nervous system related polynucleotide SEQ ID NO 10292.	PR	14-SEP-2000;	2000US-0232400P.
XX		PR	14-SEP-2000;	2000US-0232401P.
XX		PR	14-SEP-2000;	2000US-0233063P.
KW	Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;	PR	14-SEP-2000;	2000US-0233064P.
KW	immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;	PR	14-SEP-2000;	2000US-0233065P.
KW	antiparkinsonian; antisclerol; antianemic; antiarthritic; cancer;	PR	21-SEP-2000;	2000US-0234223P.
KW	antirheumatic; hepatotropic; cerebroprotective; antinflammatory;	PR	21-SEP-2000;	2000US-0234274P.
KW	antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;	PR	25-SEP-2000;	2000US-0234997P.
KW	antiparasitic; cardiant; immune disorder; cardiovascular disorder;	PR	25-SEP-2000;	2000US-0234998P.
KW	neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.	PR	26-SEP-2000;	2000US-0235484P.
XX		PR	27-SEP-2000;	2000US-0235834P.
OS	Homo sapiens.	PR	27-SEP-2000;	2000US-0235836P.
XX		PR	29-SEP-2000;	2000US-0236327P.
PN	WO200159063-A2.	PR	29-SEP-2000;	2000US-0236367P.
XX		PR	29-SEP-2000;	2000US-0236368P.
PD	16-AUG-2001.	PR	29-SEP-2000;	2000US-0236369P.
XX		PR	29-SEP-2000;	2000US-0236370P.
XX		PR	02-OCT-2000;	2000US-0236802P.
XX		PR	02-OCT-2000;	2000US-0237037P.
XX		PR	02-OCT-2000;	2000US-0237038P.
PR	31-JAN-2000;	PR	02-OCT-2000;	2000US-0237039P.
PR	04-FEB-2000;	PR	02-OCT-2000;	2000US-0237040P.
PR	24-FEB-2000;	PR	13-OCT-2000;	2000US-0239335P.
PR	02-MAR-2000;	PR	13-OCT-2000;	2000US-0239337P.
PR	16-MAR-2000;	PR	20-OCT-2000;	2000US-0240960P.
PR	17-MAR-2000;	PR	20-OCT-2000;	2000US-0241785P.
PR	18-APR-2000;	PR	20-OCT-2000;	2000US-0241786P.
PR	19-MAY-2000;	PR	20-OCT-2000;	2000US-0241787P.
PR	07-JUN-2000;	PR	20-OCT-2000;	2000US-0241808P.
PR	28-JUN-2000;	PR	20-OCT-2000;	2000US-0241809P.
PR	30-JUN-2000;	PR	20-OCT-2000;	2000US-0241826P.
PR	07-JUL-2000;	PR	20-OCT-2000;	2000US-0242221P.
PR	07-JUL-2000;	PR	01-NOV-2000;	2000US-0244617P.
PR	11-JUL-2000;	PR	08-NOV-2000;	2000US-0246475P.
PR	11-JUL-2000;	PR	08-NOV-2000;	2000US-0246476P.
PR	14-JUL-2000;	PR	08-NOV-2000;	2000US-0246477P.
PR	26-JUL-2000;	PR	08-NOV-2000;	2000US-0246478P.
PR	26-JUL-2000;	PR	08-NOV-2000;	2000US-0246523P.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246533P.
PR	18-AUG-2000;	PR	08-NOV-2000;	2000US-0246610P.
PR	22-AUG-2000;	PR	08-NOV-2000;	2000US-0246611P.
PR	22-AUG-2000;	PR	08-NOV-2000;	2000US-0246613P.
PR	22-AUG-2000;	PR	17-NOV-2000;	2000US-0249207P.
PR	23-AUG-2000;	PR	17-NOV-2000;	2000US-0249208P.
PR	30-AUG-2000;	PR	17-NOV-2000;	2000US-0249209P.
PR	01-SEP-2000;	PR	17-NOV-2000;	2000US-0249210P.
PR	01-SEP-2000;	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	PR	17-NOV-2000;	2000US-0249212P.
PR	05-SEP-2000;	PR	17-NOV-2000;	2000US-0249213P.
PR	05-SEP-2000;	PR	17-NOV-2000;	2000US-0249214P.
PR	06-SEP-2000;	PR	17-NOV-2000;	2000US-0249215P.
PR	06-SEP-2000;	PR	17-NOV-2000;	2000US-0249216P.
PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249217P.
PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249218P.
PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249300P.
PR	08-SEP-2000;	PR	01-DEC-2000;	2000US-0250391P.



PR 25-SEP-2000; 2000US-0234997P.  
 PR 26-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0234999P.  
 PR 27-SEP-2000; 2000US-0235000P.  
 PR 27-SEP-2000; 2000US-0235001P.  
 PR 27-SEP-2000; 2000US-0235002P.  
 PR 27-SEP-2000; 2000US-0235003P.  
 PR 27-SEP-2000; 2000US-0235004P.  
 PR 27-SEP-2000; 2000US-0235005P.  
 PR 27-SEP-2000; 2000US-0235006P.  
 PR 27-SEP-2000; 2000US-0235007P.  
 PR 27-SEP-2000; 2000US-0235008P.  
 PR 27-SEP-2000; 2000US-0235009P.  
 PR 27-SEP-2000; 2000US-0235010P.  
 PR 27-SEP-2000; 2000US-0235011P.  
 PR 27-SEP-2000; 2000US-0235012P.  
 PR 27-SEP-2000; 2000US-0235013P.  
 PR 27-SEP-2000; 2000US-0235014P.  
 PR 27-SEP-2000; 2000US-0235015P.  
 PR 27-SEP-2000; 2000US-0235016P.  
 PR 27-SEP-2000; 2000US-0235017P.  
 PR 27-SEP-2000; 2000US-0235018P.  
 PR 27-SEP-2000; 2000US-0235019P.  
 PR 27-SEP-2000; 2000US-0235020P.  
 PR 27-SEP-2000; 2000US-0235021P.  
 PR 27-SEP-2000; 2000US-0235022P.  
 PR 27-SEP-2000; 2000US-0235023P.  
 PR 27-SEP-2000; 2000US-0235024P.  
 PR 27-SEP-2000; 2000US-0235025P.  
 PR 27-SEP-2000; 2000US-0235026P.  
 PR 27-SEP-2000; 2000US-0235027P.  
 PR 27-SEP-2000; 2000US-0235028P.  
 PR 27-SEP-2000; 2000US-0235029P.  
 PR 27-SEP-2000; 2000US-0235030P.  
 PR 27-SEP-2000; 2000US-0235031P.  
 PR 27-SEP-2000; 2000US-0235032P.  
 PR 27-SEP-2000; 2000US-0235033P.  
 PR 27-SEP-2000; 2000US-0235034P.  
 PR 27-SEP-2000; 2000US-0235035P.  
 PR 27-SEP-2000; 2000US-0235036P.  
 PR 27-SEP-2000; 2000US-0235037P.  
 PR 27-SEP-2000; 2000US-0235038P.  
 PR 27-SEP-2000; 2000US-0235039P.  
 PR 27-SEP-2000; 2000US-0235040P.  
 PR 27-SEP-2000; 2000US-0235041P.  
 PR 27-SEP-2000; 2000US-0235042P.  
 PR 27-SEP-2000; 2000US-0235043P.  
 PR 27-SEP-2000; 2000US-0235044P.  
 PR 27-SEP-2000; 2000US-0235045P.  
 PR 27-SEP-2000; 2000US-0235046P.  
 PR 27-SEP-2000; 2000US-0235047P.  
 PR 27-SEP-2000; 2000US-0235048P.  
 PR 27-SEP-2000; 2000US-0235049P.  
 PR 27-SEP-2000; 2000US-0235050P.  
 PR 27-SEP-2000; 2000US-0235051P.  
 PR 27-SEP-2000; 2000US-0235052P.  
 PR 27-SEP-2000; 2000US-0235053P.  
 PR 27-SEP-2000; 2000US-0235054P.  
 PR 27-SEP-2000; 2000US-0235055P.  
 PR 27-SEP-2000; 2000US-0235056P.  
 PR 27-SEP-2000; 2000US-0235057P.  
 PR 27-SEP-2000; 2000US-0235058P.  
 PR 27-SEP-2000; 2000US-0235059P.  
 PR 27-SEP-2000; 2000US-0235060P.  
 PR 27-SEP-2000; 2000US-0235061P.  
 PR 27-SEP-2000; 2000US-0235062P.  
 PR 27-SEP-2000; 2000US-0235063P.  
 PR 27-SEP-2000; 2000US-0235064P.  
 PR 27-SEP-2000; 2000US-0235065P.  
 PR 27-SEP-2000; 2000US-0235066P.  
 PR 27-SEP-2000; 2000US-0235067P.  
 PR 27-SEP-2000; 2000US-0235068P.  
 PR 27-SEP-2000; 2000US-0235069P.  
 PR 27-SEP-2000; 2000US-0235070P.  
 PR 27-SEP-2000; 2000US-0235071P.  
 PR 27-SEP-2000; 2000US-0235072P.  
 PR 27-SEP-2000; 2000US-0235073P.  
 PR 27-SEP-2000; 2000US-0235074P.  
 PR 27-SEP-2000; 2000US-0235075P.  
 PR 27-SEP-2000; 2000US-0235076P.  
 PR 27-SEP-2000; 2000US-0235077P.  
 PR 27-SEP-2000; 2000US-0235078P.  
 PR 27-SEP-2000; 2000US-0235079P.  
 PR 27-SEP-2000; 2000US-0235080P.  
 PR 27-SEP-2000; 2000US-0235081P.  
 PR 27-SEP-2000; 2000US-0235082P.  
 PR 27-SEP-2000; 2000US-0235083P.  
 PR 27-SEP-2000; 2000US-0235084P.  
 PR 27-SEP-2000; 2000US-0235085P.  
 PR 27-SEP-2000; 2000US-0235086P.  
 PR 27-SEP-2000; 2000US-0235087P.  
 PR 27-SEP-2000; 2000US-0235088P.  
 PR 27-SEP-2000; 2000US-0235089P.  
 PR 27-SEP-2000; 2000US-0235090P.  
 PR 27-SEP-2000; 2000US-0235091P.  
 PR 27-SEP-2000; 2000US-0235092P.  
 PR 27-SEP-2000; 2000US-0235093P.  
 PR 27-SEP-2000; 2000US-0235094P.  
 PR 27-SEP-2000; 2000US-0235095P.  
 PR 27-SEP-2000; 2000US-0235096P.  
 PR 27-SEP-2000; 2000US-0235097P.  
 PR 27-SEP-2000; 2000US-0235098P.  
 PR 27-SEP-2000; 2000US-0235099P.  
 PR 27-SEP-2000; 2000US-0235100P.

PR 05-JAN-2001; 2001US-0259678P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-541565/60.  
 XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating nervous system cancers  
 PT and metastases.  
 XX Disclosure; SEQ ID NO 10294; 1701pp + Sequence Listing; English.  
 XX The invention relates to novel genes (ABA11004-ABA21534) and proteins  
 CC (ABA14678-ABA18001) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 32192 BP; 7965 A; 7802 C; 8055 G; 8370 T; 0 U; 0 Other;

Query Match 17.1%; Score 31.6; DB 5; Length 32192;  
 Best Local Similarity 53.8%; Pred. No. 24;  
 Matches 64; Conservative 0; Mismatches 55; Indels 0; Gaps 0;  
 QY 56 AGAGACATCCATGCCGGTTAAAGCAGATATCGTTCCATTAACAGAGAGAGGCTGCATGA 115  
 Db 14550 AGAGGTTTAATGGGCTCAAAAAGTGGCTGCTCTCCACAGAGAAAGCTCTAAAA 14491  
 QY 116 AAGGAGGTGATGGTGTTCATCTTAGGATGACAGAAATACGGATGAAAAAGAG 174  
 Db 14490 TAGAACGTGGCAGGAATACATTTAAGAGTGGGAGAGAGTAACACAGATAAGAGGAG 14432

RESULT 8  
 ADH69807\_4/c  
 Continuation (5 of 7) of ADH69807 from base 400001 (Human Vbeta gene.)  
 WP Sequence split into 7 fragments LOCUS ADH69807 Accession Adh69807  
 WP Fragment Name Begin End  
 WP ADH69807\_0 110000  
 WP ADH69807\_1 100001 210000  
 WP ADH69807\_2 200001 310000  
 WP ADH69807\_3 300001 410000  
 WP ADH69807\_4 400001 510000  
 WP ADH69807\_5 500001 610000  
 WP ADH69807\_6 600001 684973

Query Match 17.1%; Score 31.6; DB 12; Length 110000;  
 Best Local Similarity 49.1%; Pred. No. 36;  
 Matches 82; Conservative 0; Mismatches 85; Indels 0; Gaps 0;  
 QY 19 TCTATTTTCATTGATCAAGGATTTATACACAAACAGAGAGATCATCGCGGTTTAAA 78  
 Db 104161 TCCATTGCAACTGTTTGTAAAGCTATGAACATATAAACTACATAGCTTTGAGTGCCTGTA 104102  
 QY 79 GCAGTATCGTTCCATCTTAACAGAGAGAGGCTGATGAAGAGAGGTGATGGTTTTCATC 138  
 Db 104101 ATAGCAAGATTTTACCTTAATTAAGAGTTGGGAAGAGAGTGGCATTTGAGTTGGTTC 104042  
 QY 139 TTAGGGATGACAGACAATACGGATGAAAAAGGAGGAGGATGAAA 185

Db 104041 TGAAGGATGACAGAGAACTGCTGCAACCCAGAGAGAGAAACA 103995

RESULT 9  
ADH69807\_5/c  
Continuation (6 of 7) of ADH69807 from base 500001 (Human Vbeta gene.)  
WP Sequence split into 7 fragments LOCUS ADH69807 Accession Adh69807

Fragment Name	Begin	End
WP ADH69807_0	1	110000
WP ADH69807_1	100001	210000
WP ADH69807_2	200001	310000
WP ADH69807_3	300001	410000
WP ADH69807_4	400001	510000
WP ADH69807_5	500001	610000
WP ADH69807_6	600001	684973

Query Match 17.1%; Score 31.6; DB 12; Length 110000;  
Best Local Similarity 49.1%; Pred. No. 36;  
Matches 82; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 19 TCTATTTCATTGTATCAAGAGTTTATACAAACAGAGACATCCATGCCGGGTAAA 78  
DB 4161 TCCATTGCAACTGTTTGTAAAGCTATGAACATAAAAACTACATAGCTTTGAGTGCCTGTA 4102  
QY 79 GCAGTATCGTTCATCTACAGAGAGGCTCATGAAAGGAGGTGATGGGTTTTTCATC 138  
DB 4101 ATAGCAGATTTTACCTATTAATAAGGTTGGAAAGGAGTGGCATTTGAGTTGGTTTC 4042  
QY 139 TTAGGATGACAGAAACAATPACGATGAAAAAAGGAGGAGATGGAAA 185  
DB 4041 TGAAGGATGACAGAGAACTGCTGCAACCCAGAGAGAGAAACA 3995

RESULT 10  
AAV21209\_04  
Continuation (5 of 17) of AAV21209 from base 400001 (Methanococcus jannaschii circular  
WP Sequence split into 17 fragments LOCUS AAV21209 Accession Aav21209

Fragment Name	Begin	End
WP AAV21209_00	1	110000
WP AAV21209_01	100001	210000
WP AAV21209_02	200001	310000
WP AAV21209_03	300001	410000
WP AAV21209_04	400001	510000
WP AAV21209_05	500001	610000
WP AAV21209_06	600001	710000
WP AAV21209_07	700001	810000
WP AAV21209_08	800001	910000
WP AAV21209_09	900001	1010000
WP AAV21209_10	1000001	1110000
WP AAV21209_11	1100001	1210000
WP AAV21209_12	1200001	1310000
WP AAV21209_13	1300001	1410000
WP AAV21209_14	1400001	1510000
WP AAV21209_15	1500001	1610000
WP AAV21209_16	1600001	1664976

Query Match 17.0%; Score 31.4; DB 2; Length 110000;  
Best Local Similarity 59.6%; Pred. No. 42;  
Matches 53; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 17 GATCTATTTTCATTGTATCAAGGATTTATACAAACAGAGACATCCATGCCGGGTAA 76  
DB 92343 GAATCATGTCGATTTGAATCGTAGTTTATCATCTTTAAGCTACTTACAGCCCTTCTA 92402  
QY 77 AAGCAGTATCGTTTCATCTAACAGAGAAG 105  
DB 92403 ATTCAATTTTTTCTTTTCACTGAGAAG 92431

RESULT 11  
AAI91410  
ID AAI91410 standard; cDNA; 443 BP.

XX AAI91410;  
XX AC  
DT 06-NOV-2001 (first entry)  
XX Human polynucleotide SEQ ID NO 11470.  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation; ss.  
OS Homo sapiens.  
PN WO200164835-A2.  
XX WO200164835-A2.  
PD 07-SEP-2001.  
XX 26-FEB-2001; 2001WO-US004927.  
XX 28-FEB-2000; 2000US-00515126.  
PR 18-MAY-2000; 2000US-00577409.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT;  
XX WPI; 2001-514838/56.  
DR P-PSDB; AAO11479.  
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
PT and treating e.g. leukemia, inflammation and immune disorders.  
PS Claim 1; SEQ ID NO 11470; 1399pp + Sequence Listing; English.  
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 443 BP; 143 A; 80 C; 112 G; 108 T; 0 U; 0 Other;  
SQ  
Query Match 16.9%; Score 31.2; DB 4; Length 443;  
Best Local Similarity 50.3%; Pred. No. 7.6;  
Matches 75; Conservative 0; Mismatches 74; Indels 0; Gaps 0;  
QY 37 AAGGATTTATACAAACAGAGACATCCATGCCGGTTAAAGCAGATATCGTTCATCTA 96  
DB 132 AAGGATAGCAAGCGCAGACAAACATAGATAATTTTCAAGGAGTTCACTCTAATG 191  
QY 97 ACAGAGAAGNCTGATGAAGAGGAGGTGATGGGTTTTTCATCTTAGGATGACAGAACAA 156  
DB 192 GGAGGAAAGACAAGCGAAGAGGTGAAGAGAGACAGGAAGCTAGGGCAGTCAGAAATG 251  
QY 157 TACGGATGAAAAAGGAGAGGGATGGAAA 185  
DB 252 GAGCAAGGGGAAAGAAAGGAATCGAGA 280  
RESULT 12  
ABK39185  
ID ABK39185 standard; cDNA; 390 BP.  
XX AC  
XX ABK39185;

XX 21-MAY-2002 (first entry)  
DT ACAL1514;  
XX CDNA encoding lung tumour protein clone R0131:G04.  
DE  
XX Lung tumour; cancer; T cell; immune response stimulator; cytostatic;  
KW gene; ss.  
XX Homo sapiens.  
OS  
XX WO200204514-A2.  
PN  
XX 17-JAN-2002.  
XX  
XX 10-JUL-2001; 2001WO-US022058.  
XX  
XX 11-JUL-2000; 2000US-00614124.  
PR 29-AUG-2000; 2000US-00651563.  
PR 08-SEP-2000; 2000US-00658824.  
PR 26-SEP-2000; 2000US-00671325.  
PR 06-OCT-2000; 2000US-00677419.  
PR 30-OCT-2000; 2000US-00702705.  
PR 13-DEC-2000; 2000US-00736457.  
PR 03-MAY-2001; 2001US-00849626.  
XX (CORI-) CORIXA CORP.  
PA  
XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;  
PI Marnerakis M, Carter D, Fanger GR, Vedwick TS, Bangur CS, Mcnabb A;  
PI Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;  
XX  
XX WPI; 2002-164634/21.  
XX  
XX Novel polynucleotide encoding a lung tumor polypeptide useful for  
PT stimulating and/or expanding T cells specific for a tumor protein.  
PT  
XX  
XX Claim 1; SEQ ID NO 1223; 223pp; English.  
PS  
XX The invention describes an isolated polynucleotide and polypeptide useful  
XX for stimulating and/or expanding T cells specific for a tumour protein  
CC for determining the presence of a cancer in a patient. A composition  
CC containing the polynucleotide and/or polypeptide is useful for treating a  
CC lung cancer in a patient. The polypeptide is useful for removing tumour  
CC cells from a biological sample. The polynucleotide is also useful as  
CC a probe or primer to detect the level of mRNA encoding a tumour protein.  
CC This sequence encodes a lung tumour associated protein or protein  
CC fragment, described in the method of the invention. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 390 BP; 96 A; 92 C; 83 G; 119 T; 0 U; 0 Other;  
SQ  
  
Query Match 16.8%; Score 31; DB 6; Length 390;  
Best Local Similarity 48.8%; Pred. No. 8.4;  
Matches 82; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
  
QY 14 CTTGATCTATTTCATTTGTATCAAGGATTATACACAAACAGACATCCATGCGGG 73  
DB 175 CATGAATCTCTGTTCTGACGAGTGTCTCACAGTGTCTGACGTCCTCGACTCGAG 234  
  
QY 74 TTAAGCAGTATCTGTTCCATCTACAGAGAGGCTCATGAAGGAGGTGATGGTTTT 133  
DB 235 TTTATTGGTGTGTTTCTTCTTGAGATCCATGCAATTTCTGGTTGATCTCTGGAATC 294  
  
QY 134 TCATCTTAGGATGCACAGACAAATACCGATGAAAAAGGAGGAGTG 181  
DB 295 CCTATTAGGTATGAATAGCATGATGATGATGATGATGATGATGATGATGATGATG 342  
  
RESULT 13  
ACAL1514  
ID ACAL1514 standard; cDNA; 390 BP.

XX ACAL1514;  
AC  
XX 05-JUN-2003 (first entry)  
DT  
XX Human lung adenocarcinoma library cDNA SEQ ID 1223.  
DE  
XX Human; lung cancer; ss; lung tumour; cytostatic; vaccine;  
KW T cell expansion; CD4; CD8.  
XX  
XX Homo sapiens.  
OS  
XX US2002197669-A1.  
PN  
XX 26-DEC-2002.  
XX  
XX 03-MAY-2001; 2001US-00849626.  
XX  
XX 13-DEC-2000; 2000US-00736457.  
XX (BANG/) BANGUR C S.  
PA (WANG/) FANGER G R.  
PA (WANG/) WANG A.  
PA (WANG/) WANG T.  
PA (SWIT/) SWITZER A P.  
PA (MCNE/) MCNEILL P D.  
PA (CLAP/) CLAPPER J D.  
XX  
XX Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;  
PI Clapper JD;  
PI  
XX WPI; 2003-352750/33.  
XX  
XX Novel lung cancer polynucleotide encoding lung cancer protein, useful for  
PT detecting the presence of lung cancer in a patient, and in pharmaceutical  
PT compositions, e.g. vaccines, for treating lung cancer.  
XX  
XX Example 1; Page; 72pp; English.  
PS  
XX The invention relates to a polynucleotide encoding a lung tumour protein,  
CC comprising a sequence selected from any of the 14 sequences mentioned in  
CC the specification, or a sequence (S2) mentioned in specification,  
CC complement of S1, sequences consisting of at least 20 contiguous residues  
CC of S1, sequences that hybridise to S1, sequences having 75%, preferably  
CC 90%, identity to S1, or degenerate variants of S1. Also included are an  
CC isolated polypeptide (comprising a sequence (S3) selected from any one of  
CC the 4 amino acid sequences mentioned in the specification, a sequence  
CC encoded by the polynucleotide, or a sequence encoded by the polynucleotide), an  
CC preferably 90%, identity to a sequence encoded by the polynucleotide, an  
CC expression vector comprising the polynucleotide operably linked to an  
CC expression control sequence, a host cell transformed or transfected with  
CC the vector, an isolated antibody (or its antigen-binding fragment) that  
CC specifically binds to the polypeptide, detecting the presence of a cancer  
CC in a patient, a fusion protein comprising the polypeptide, an  
CC oligonucleotide that hybridises to S1 under moderately stringent  
CC conditions, stimulating and/or expanding T cells specific for a tumour  
CC protein (comprising contacting T cells with the polynucleotide, protein  
CC or antigen-presenting cells, under conditions and for a time sufficient  
CC to permit the stimulation and/or expansion of T cells) and inhibiting the  
CC development of a cancer in a patient (by incubating CD4+ and/or CD8+ T  
CC cells isolated from a patient with the polynucleotide, protein or antigen  
CC presenting cells that express the polynucleotide, such that T cells  
CC proliferate, administering to the patient an effective amount of the  
CC proliferated T cells, and thus inhibiting the development of a cancer in  
CC the patient. The polynucleotide, protein and cells are useful in a  
CC composition for stimulating an immune response in a patient, and for  
CC treating a cancer in a patient (particularly lung cancer). The  
CC oligonucleotide is useful for determining the presence of a cancer in a  
CC patient. The protein and oligonucleotides are useful in pharmaceutical  
CC compositions, e.g. vaccines. The polynucleotide is also useful as a probe  
CC or primer for nucleic acid hybridisation, and in the design and  
CC preparation of ribozyme molecules for inhibiting expression of tumour  
CC polypeptides and proteins in tumour cells. An amplified portion of the

CC polynucleotide is useful for isolating a full-length gene from a suitable  
 CC library. The present sequence is a cDNA (full length, extended or  
 CC partial) isolated from a library derived from lung tumour/cancer cells.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC USPTO at seqdata.uspto.gov/sequence.html?DocId=20020197669

XX Sequence 390 BP; 96 A; 92 C; 83 G; 119 T; 0 U; 0 Other;

Query Match 16.8%; Score 31; DB 8; Length 390;  
 Best Local Similarity 48.8%; Pred. No. 8.4;  
 Matches 82; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 14 CTTGATCTATTTTCATTGTATCAAGGATTTATACAAAGAGACATCCATCCCGGG 73  
 Db 175 CATGAATCTCTGTTTCGTAGCAAGTGCATGCTCACAGTTGTCAGTCTGCCACTCCGAG 234  
 QY 74 TTAAAGCAGTATCGTTCCATCTAACAGAGAAGNCTGCATGAAAGAGAGGATG 181  
 Db 235 TTTATTGGTGTGTTTCTCTTGAGATCCATGCAATTCCTGTTGAAATCTCTCGGAATC 294  
 QY 134 TCATCTTAGGATGACAGACAAATACGATGAAAGAGAGGATG 181  
 Db 295 CCTCATTAGGTATGAATAGCATGATGCAATTCATATAAAGTCACGAAGG 342

RESULT 14  
 ACA02700  
 ID ACA02700 standard; cDNA; 390 BP.

XX ACA02700;  
 AC (first entry)  
 DT 22-MAY-2003 (first entry)

XX Lung cancer therapy and diagnosis associated cDNA #1189.

DE Lung cancer; cytostatic; vaccine; gene therapy; cancer; gene; ss.

XX Homo sapiens.

XX US2002172952-A1.

PN 21-NOV-2002.

XX 10-JUL-2001; 2001US-00902941.

XX 30-JUN-1999; 99US-00346492.

PR 15-OCT-1999; 99US-00419356.

PR 17-DEC-1999; 99US-00466867.

PR 30-DEC-1999; 99US-00476300.

PR 06-MAR-2000; 2000US-00519642.

PR 22-MAR-2000; 2000US-00533077.

PR 10-APR-2000; 2000US-00546259.

PR 27-APR-2000; 2000US-00560406.

PR 05-JUN-2000; 2000US-00589184.

PR 11-JUL-2000; 2000US-00614124.

PR 29-AUG-2000; 2000US-00651563.

PR 08-SEP-2000; 2000US-00658824.

PR 26-SEP-2000; 2000US-00671325.

PR 06-OCT-2000; 2000US-00677419.

PR 30-OCT-2000; 2000US-00702705.

PR 13-DEC-2000; 2000US-00736457.

PR 03-MAY-2001; 2001US-00849626.

XX (CORI-) CORIXA CORP.

XX Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW;

PI Durham M, Carter D, Fanger GR, Vedwick TS, Bangur CS, McNabb A;

XX WPI; 2003-328427/31.

XX New polynucleotide, useful for preparing a composition for treating or

PT inhibiting development of cancer, e.g. lung cancer.

XX Example 1; SEQ ID NO 1223; 82pp; English.

XX The invention describes an isolated polynucleotide comprising one of 32  
 CC sequences, complement or degenerate variants of them. The polynucleotide  
 CC is useful for preparing a composition e.g. a vaccine or for gene therapy,  
 CC for treating or inhibiting development of cancer, e.g. lung cancer. This  
 CC sequence represents a polynucleotide associated with the compositions and  
 CC methods for the therapy and diagnosis of lung cancer

XX Sequence 390 BP; 96 A; 92 C; 83 G; 119 T; 0 U; 0 Other;

Query Match 16.8%; Score 31; DB 8; Length 390;  
 Best Local Similarity 48.8%; Pred. No. 8.4;  
 Matches 82; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 14 CTTGATCTATTTTCATTGTATCAAGGATTTATACAAAGAGACATCCATCCCGGG 73  
 Db 175 CATGAATCTCTGTTTCGTAGCAAGTGCATGCTCACAGTTGTCAGTCTGCCACTCCGAG 234  
 QY 74 TTAAAGCAGTATCGTTCCATCTAACAGAGAAGNCTGCATGAAAGAGGATGATGGT 133  
 Db 235 TTTATTGGTGTGTTTCTCTTGAGATCCATGCAATTCCTGTTGAAATCTCTCGGAATC 294  
 QY 134 TCATCTTAGGATGACAGACAAATACGATGAAAGAGAGGATG 181  
 Db 295 CCTCATTAGGTATGAATAGCATGATGCAATTCATATAAAGTCACGAAGG 342

RESULT 15  
 ADH46742  
 ID ADH46742 standard; cDNA; 390 BP.

XX ADH46742;

XX 25-MAR-2004 (first entry)

DE Human lung tumour cDNA clone, SEQ ID NO 1223.

KW lung cancer; tumour; cytostatic; immunostimulant; vaccine; gene therapy;  
 human; clone; ss.

XX Homo sapiens.

PN WO2003037267-A2.

XX 08-MAY-2003.

XX 28-OCT-2002; 2002WO-US034777.

XX 29-OCT-2001; 2001US-00017754.

XX 28-MAR-2002; 2002US-00113872.

XX (CORI-) CORIXA CORP.

XX Henderson RA, Wang T, Watanabe Y, Kalos MD, Sleath PR;

PI Johnson JC, Retter MW, Durham M, Carter D, Fanger GR, Vedwick TS;

PI Bangur CS, McNabb A;

XX WPI; 2003-468346/44.

XX New polypeptides and encoding polynucleotides, useful for diagnosing,  
 PT preventing and/or treating lung cancer.  
 XX Example 1; SEQ ID NO 1223; 258pp; English.

XX The invention relates to novel compositions and methods for the therapy

CC and diagnosis of cancer, particularly lung cancer. The compositions  
 CC comprise one or more lung tumour polypeptides, immunogenic portions  
 CC thereof, polynucleotides that encode such polypeptides, antigen presenting  
 CC cells that express such polypeptides, and T cells that are specific for  
 CC cells expressing such polypeptides. The novel compositions have  
 CC cytostatic and immunostimulant activity. The lung tumour antigens can be



CC used in the creation of a vaccine. The polynucleotides that encode the  
CC lung tumour polypeptides can be used in gene therapy to help in the  
CC treatment of lung tumours. This polynucleotide sequence represents a  
CC human lung tumour cDNA clone of the invention. This sequence was not  
CC shown in the specification. It has been taken from a World Intellectual  
CC Property Organization CD ROM supplied with the specification.  
XX  
SQ Sequence 390 BP; 96 A; 92 C; 83 G; 119 T; 0 U; 0 Other;  
Query Match 16.8%; Score 31; DB 10; Length 390;  
Best Local Similarity 48.8%; Pred. NO. 8.4;  
Matches 82; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
QY 14 CTTGATCTATTTCATTTGATCAAGGATTATACACAAACAAGAGACATCCATGCCGG 73  
DB 175 CATGAATCTCTGTTTCGTAGCAAGTGCACTCTCAGTTGTCTGCCACTCCGAG 234  
QY 74 TTAAGCAGTATCGTTCCATCTAACAGAGAAGGNTGTCATGAAGAGGATGATGGTTTT 133  
DB 235 TTTATTGGTGTGTTTCTTTGAGATCCATGCTTCTGTTGAATCTCCTGGAACTC 294  
QY 134 TCATCTTAGGGATGACAGAACATACGGATGAAAAAGGAGAGGATG 181  
DB 295 CCTCATTAGGTATGAATAGCATGATGTCATTGCATTAAGTCACGAAGG 342  
Search completed: November 10, 2004, 17:09:37  
Job time : 121.352 secs

**THIS PAGE BLANK (USPTO)**



```

Query Match          99.5%; Score 184; DB 3; Length 185;
Best Local Similarity 100.0%; Pred. No. 5.7e-52; Indels 0; Gaps 0;
Matches 185; Conservative 0; Mismatches 0;

QY 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAGAGATTATACACAAACAAGAGA 60
   |||||
DB 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAGAGATTATACACAAACAAGAGA 60

QY 61 CATCCATGCCGGGTTAAAGCAGTATCGTTCATCTTAACAGAGAAGGCTGCATGAAAGGA 120
   |||||
DB 61 CATCCATGCCGGGTTAAAGCAGTATCGTTCATCTTAACAGAGAAGGCTGCATGAAAGGA 120

QY 121 GGTGATGGGTTTTTCATCTTACGGGATGACAGAACATACGGATGAAAAAGGAGAGGAT 180
   |||||
DB 121 GGTGATGGGTTTTTCATCTTACGGGATGACAGAACATACGGATGAAAAAGGAGAGGAT 180

QY 181 GGAAA 185
   |||||
DB 181 GGAAA 185

RESULT 2
US-08-470-953A-26
; Sequence 26, Application US/08470953A
; Patent No. 6346407
; GENERAL INFORMATION:
; APPLICANT: ANDREE LAHAYE
; APPLICANT: ERIC DE BUYL
; APPLICANT: PIERRE LEDOUX
; APPLICANT: RENE DETROZ
; TITLE OF INVENTION: Xylanase, microorganisms produced it,
; TITLE OF INVENTION: DNA molecule, processes for preparation of this xylanase
; TITLE OF INVENTION: and uses thereof
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE
; STREET: 2000 K St., N.W., Suite 200
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,953A
; FILING DATE: 6-OCTOBER-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilhem P. Gadiano, Esq.
; REGISTRATION NUMBER: 37,136
; REFERENCE/DOCKET NUMBER: 4121-40
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-429-0625
; TELEFAX: (202) 293-1850
; TELEX: 650 383-5605
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 185 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-470-953A-26

Query Match          99.5%; Score 184; DB 3; Length 185;
Best Local Similarity 100.0%; Pred. No. 5.7e-52; Indels 0; Gaps 0;
Matches 185; Conservative 0; Mismatches 0;

```

```

QY 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAGAGATTATACACAAACAAGAGA 60
   |||||
DB 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGTATCAAGAGATTATACACAAACAAGAGA 60

QY 61 CATCCATGCCGGGTTAAAGCAGTATCGTTCATCTTAACAGAGAAGGCTGCATGAAAGGA 120
   |||||
DB 61 CATCCATGCCGGGTTAAAGCAGTATCGTTCATCTTAACAGAGAAGGCTGCATGAAAGGA 120

QY 121 GGTGATGGGTTTTTCATCTTACGGGATGACAGAACATACGGATGAAAAAGGAGAGGAT 180
   |||||
DB 121 GGTGATGGGTTTTTCATCTTACGGGATGACAGAACATACGGATGAAAAAGGAGAGGAT 180

QY 181 GGAAA 185
   |||||
DB 181 GGAAA 185

RESULT 3
US-09-076-677-28
; Sequence 28, Application US/09076677
; Patent No. 6423523
; GENERAL INFORMATION:
; APPLICANT: DE BUYL, ERIC
; APPLICANT: LAHAYE, ANDREE
; APPLICANT: LEDOUX, PIERRE
; APPLICANT: AMORY, ANTOINE
; APPLICANT: DETROZ, RENE
; APPLICANT: ANDRE, CHRISTOPHE
; APPLICANT: VETTER, ROMAN
; TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,
; EXPRESSION VECTORS FOR SUCH XYLANASE AND
; OTHER PROTEINS, HOST ORGANISMS THEREFOR AND
; USE THEREOF
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.
; STREET: 2000 K St., N.W., Suite 200
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/076,677
; FILING DATE: 12-May-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/275,526
; FILING DATE: 15-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Gadiano, Wilhem F.
; REGISTRATION NUMBER: 37,136
; REFERENCE/DOCKET NUMBER: 4121-49
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 429-0625
; TELEFAX: (202) 293-0625
; TELEX: 650 383 5605
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 185 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-076-677-28

Query Match          99.5%; Score 184; DB 4; Length 185;

```

	Best Local Similarity	100.08;	Pred. No. 5.7e-52;		
	Matches 185;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
1	TCATGTAAC	TCGCCCTTGATCTATTTTCATTTGGTATCAAAAGGATTTATATACACAAACAGAGA	60		
1	TCATGTAAC	TCGCCCTTGATCTATTTTCATTTGGTATCAAAAGGATTTATATACACAAACAGAGA	60		
61	CATCCATCGCGGTTAAAGCAGTATCGTTCCATCTAAACAGAGAAGGNTGTCATGAAAGGA	120			
61	CATCCATCGCGGTTAAAGCAGTATCGTTCCATCTAAACAGAGAAGGNTGTCATGAAAGGA	120			
121	GGTATCGGTTTTTCATCTTAGGGATGACAGACAAATACGGATGAAAAAGGAGAGGGAT	180			
121	GGTATCGGTTTTTCATCTTAGGGATGACAGACAAATACGGATGAAAAAGGAGAGGGAT	180			
181	GGAAA	185			
181	GGAAA	185			

## RESULT 4

US-09-073-055-28  
; Sequence 28, Application US/09073055  
; Patent No. 6426211  
; GENERAL INFORMATION:  
; APPLICANT: DE BUYL, ERIC  
; LAHAYE, ANDR E  
; LEDOUX, PIERRE  
; AMORY, ANTOINE  
; DETROZ, REN  
; ANDRE, CHRISTOPHE  
; VETTER, ROMAN  
; TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,  
; EXPRESSION VECTORS FOR SUCH XYLANASE AND  
; OTHER PROTEINS, HOST ORGANISMS THEREFOR AND  
; USE THEREOF

NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.  
STREET: 2000 K St., N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patencin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/073,055  
FILING DATE: 05-May-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/275,526  
FILING DATE: 15-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Gadiano, Wilhem F.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-49  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 429-0625  
TELEFAX: (202) 293-0625  
TELEX: 650 383 5605  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 185 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 28:  
US-09-073-055-28

US-09-073-055-28

	Query Match	99.5%;	Score 184;	DB 4;	Length 185;
	Best Local Similarity	100.0%;	Pred. No. 5.7e-52;		
	Matches 185;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
Qy	1	TCATGTAACTCGCGCTTGATCTATTTTCATTGTGTATCAAAGGATTTATACAAACAAGAGA	60		
Dd	1	TCATGTAACTCGCGCTTGATCTATTTTCATTGTGTATCAAAGGATTTATACAAACAAGAGA	60		
Qy	61	CATCCATGCGGGGTTAAAGCAGTATCGTTCCATCTTACAGAGAAGGCTGCATGAAAGGA	120		
Dd	61	CATCCATGCGGGGTTAAAGCAGTATCGTTCCATCTTACAGAGAAGGCTGCATGAAAGGA	120		
Qy	121	GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACGGATCGAAAAAAGGAGAGGGAT	180		
Dd	121	GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACGGATCGAAAAAAGGAGAGGGAT	180		
Qy	181	GGAAA	185		
Dd	181	GGAAA	185		

## RESULT 5

US-08-275-526C-1  
 Sequence 1, Application US/08275526C  
 Patent No. 6180382  
 GENERAL INFORMATION:  
 APPLICANT: DE BUYL, ERIC  
 APPLICANT: LAHAYE, ANDR E  
 APPLICANT: LEDOUX, PIERRE  
 APPLICANT: AMORY, ANTOINE  
 APPLICANT: DETROZ, REN  
 APPLICANT: ANDRE, CHRISTOPHE  
 APPLICANT: VETTER, ROMAN  
 TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,  
 TITLE OF INVENTION: EXPRESSION VECTORS FOR SUCH XYLANASE AND  
 TITLE OF INVENTION: OTHER PROTEINS, HOST ORGANISMS THEREFOR AND  
 TITLE OF INVENTION: USE THEREOF  
 NUMBER OF SEQUENCES: 35  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.  
 STREET: 2000 K St., N.W., Suite 200  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20006  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC Compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/275,526C  
 FILING DATE: 15-JUL-1994  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Gadiano, Wilhem F.  
 REGISTRATION NUMBER: 37,136  
 REFERENCE/DOCKET NUMBER: 4121-49  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 429-0625  
 TELEFAX: (202) 293-0625  
 TELEX: 650 383 5605  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1022 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: genomic DNA  
 ORIGINAL SOURCE:  
 ORGANISM: Bacillus pumilus  
 STRAIN: PRL B12

## US-08-275-526C-1

Query Match 99.5%; Score 184; DB 3; Length 1022;  
Best Local Similarity 100.0%; Pred. No. 1.2e-51;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGATCAAGAGGATTTATACACAAACAAGAGA 60  
Db 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGATCAAGAGGATTTATACACAAACAAGAGA 60  
Qy 61 CATCCATCCCGGGTTAAAGCAGTATCGTTCATCTAAACAGAGAGGCTGCATGAAAGGA 120  
Db 61 CATCCATCCCGGGTTAAAGCAGTATCGTTCATCTAAACAGAGAGGCTGCATGAAAGGA 120  
Qy 121 GGTGATGGGTTTTTCATCTTAGGATGACAGAAACATACCGATGAAAAAGGAGGGAT 180  
Db 121 GGTGATGGGTTTTTCATCTTAGGATGACAGAAACATACCGATGAAAAAGGAGGGAT 180  
Qy 181 GGAAA 185  
Db 181 GGAAA 185

## RESULT 6

US-08-275-526C-35  
; Sequence 35, Application US/08275526C  
; Patent No. 6180382

## GENERAL INFORMATION:

APPLICANT: DE BUYL, ERIC  
APPLICANT: LAHAYE, ANDR E  
APPLICANT: LEDOUX, PIERRE  
APPLICANT: AMORY, ANTOINE  
APPLICANT: DETROZ, REN  
APPLICANT: ANDRE, CHRISTOPHE  
APPLICANT: VETTER, ROMAN

TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,  
EXPRESSION VECTORS FOR SUCH XYLANASE AND  
TITLE OF INVENTION: OTHER PROTEINS, HOST ORGANISMS THEREFOR AND  
TITLE OF INVENTION: USE THEREOF

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.

STREET: 2000 K St., N.W., Suite 200

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/275,526C

FILING DATE: 15-JUL-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Gadiano, Wilhem F.

REGISTRATION NUMBER: 37,136

REFERENCE/DOCKET NUMBER: 4121-49

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 429-0625

TELEFAX: (202) 293-0625

INFORMATION FOR SEQ ID NO: 35:

SEQUENCE CHARACTERISTICS:

LENGTH: 1022 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: genomic DNA

US-08-275-526C-35

## Query Match

99.5%; Score 184; DB 3; Length 1022;  
Best Local Similarity 100.0%; Pred. No. 1.2e-51;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGATCAAGAGGATTTATACACAAACAAGAGA 60  
Db 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGATCAAGAGGATTTATACACAAACAAGAGA 60  
Qy 61 CATCCATCCCGGGTTAAAGCAGTATCGTTCATCTAAACAGAGAGGCTGCATGAAAGGA 120  
Db 61 CATCCATCCCGGGTTAAAGCAGTATCGTTCATCTAAACAGAGAGGCTGCATGAAAGGA 120  
Qy 121 GGTGATGGGTTTTTCATCTTAGGATGACAGAAACATACCGATGAAAAAGGAGGGAT 180  
Db 121 GGTGATGGGTTTTTCATCTTAGGATGACAGAAACATACCGATGAAAAAGGAGGGAT 180  
Qy 181 GGAAA 185  
Db 181 GGAAA 185

## RESULT 7

US-09-076-677-1

; Sequence 1, Application US/09076677

; Patent No. 6423523

GENERAL INFORMATION:

APPLICANT: DE BUYL, ERIC

LAHAYE, ANDREE

LEDoux, PIERRE

AMORY, ANTOINE

DETROZ, RENE

ANDRE, CHRISTOPHE

VETTER, ROMAN

TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,

EXPRESSION VECTORS FOR SUCH XYLANASE AND

OTHER PROTEINS, HOST ORGANISMS THEREFOR AND

USE THEREOF

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE, P.C.

STREET: 2000 K St., N.W., Suite 200

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/076,677

FILING DATE: 12-May-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/275,526

FILING DATE: 15-JUL-1994

ATTORNEY/AGENT INFORMATION:

NAME: Gadiano, Wilhem F.

REGISTRATION NUMBER: 37,136

REFERENCE/DOCKET NUMBER: 4121-49

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 429-0625

TELEFAX: (202) 293-0625

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1022 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: genomic DNA

ORIGINAL SOURCE:

ORGANISM: Bacillus pumilus  
STRAIN: PRL B12  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-076-677-1

Query Match 99.5%; Score 184; DB 4; Length 1022;  
Best Local Similarity 100.0%; Pred. No. 1.2e-51;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGATCAAAAGGATTTATACACAAACAAGAGA 60  
DB 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGATCAAAAGGATTTATACACAAACAAGAGA 60

QY 61 CATCCATGCCGGGTTAAAGCAGTATCGTTCCATCTAACAGAGAGAGGCTGCATGAAGAGA 120  
DB 61 CATCCATGCCGGGTTAAAGCAGTATCGTTCCATCTAACAGAGAGAGGCTGCATGAAGAGA 120

QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACGGATGAAAAAAGGAGAGGGAT 180  
DB 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACGGATGAAAAAAGGAGAGGGAT 180

QY 181 GGAAG 185  
DB 181 GGAAG 185

RESULT 8  
US-09-076-677-35  
Sequence 35, Application US/09076677  
Patent No. 6423523  
GENERAL INFORMATION:  
APPLICANT: DE BUYL, ERIC  
LAHAYE, ANDREE  
LEDoux, PIERRE  
AMORY, ANTOINE  
DETROZ, RENE  
ANDRE, CHRISTOPHE  
VETTER, ROMAN  
TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,  
EXPRESSION VECTORS FOR SUCH XYLANASE AND  
OTHER PROTEINS, HOST ORGANISMS THEREFOR AND  
USE THEREOF  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAN BRINKS HOFER GILSON & LIONE, P.C.  
STREET: 2000 K St., N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/076.677  
FILING DATE: 12-May-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/275,526  
FILING DATE: 15-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Gadiano, Wilhem F.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-49  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 429-0625  
TELEFAX: (202) 293-5605  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1022 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 35:  
US-09-076-677-35

Query Match 99.5%; Score 184; DB 4; Length 1022;  
Best Local Similarity 100.0%; Pred. No. 1.2e-51;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGATCAAAAGGATTTATACACAAACAAGAGA 60  
DB 1 TCATGTAACCTCGCCTTGATCTATTTTCATTTGATCAAAAGGATTTATACACAAACAAGAGA 60

QY 61 CATCCATGCCGGGTTAAAGCAGTATCGTTCCATCTAACAGAGAGAGGCTGCATGAAGAGA 120  
DB 61 CATCCATGCCGGGTTAAAGCAGTATCGTTCCATCTAACAGAGAGAGGCTGCATGAAGAGA 120

QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACGGATGAAAAAAGGAGAGGGAT 180  
DB 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAACATACGGATGAAAAAAGGAGAGGGAT 180

QY 181 GGAAG 185  
DB 181 GGAAG 185

RESULT 9  
US-09-073-055-1  
Sequence 1, Application US/09073055  
Patent No. 6426211  
GENERAL INFORMATION:  
APPLICANT: DE BUYL, ERIC  
LAHAYE, ANDR E  
LEDoux, PIERRE  
AMORY, ANTOINE  
DETROZ, REN  
ANDRE, CHRISTOPHE  
VETTER, ROMAN  
TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,  
EXPRESSION VECTORS FOR SUCH XYLANASE AND  
OTHER PROTEINS, HOST ORGANISMS THEREFOR AND  
USE THEREOF  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WILLIAN BRINKS HOFER GILSON & LIONE, P.C.  
STREET: 2000 K St., N.W., Suite 200  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/073,055  
FILING DATE: 05-May-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/275,526  
FILING DATE: 15-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Gadiano, Wilhem F.  
REGISTRATION NUMBER: 37,136  
REFERENCE/DOCKET NUMBER: 4121-49  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 429-0625  
TELEFAX: (202) 293-5605  
INFORMATION FOR SEQ ID NO: 1:

```

/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1022 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: genomic DNA
/ ORIGINAL SOURCE:
/ ORGANISM: Bacillus pumilus
/ STRAIN: PRL B12
/ SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-073-055-1

Query Match          99.5%; Score 184; DB 4; Length 1022;
Best Local Similarity 100.0%; Pred. No. 1.2e-51;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATGTAACGCTTGTATCTATTCATTGTTATCAAGGATTTATACAAACAAAGAGA 60
Db 1 TCATGTAACGCTTGTATCTATTCATTGTTATCAAGGATTTATACAAACAAAGAGA 60
QY 61 CATCCATGCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAGGCTGCATGAAGGA 120
Db 61 CATCCATGCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAGGCTGCATGAAGGA 120
QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAAACATACGGATGAAAAAGGAGGGAT 180
Db 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAAACATACGGATGAAAAAGGAGGGAT 180
QY 181 GGAAA 185
Db 181 GGAAA 185

```

```

RESULT 10
US-09-073-055-35
/ Sequence 35, Application US/09073055
/ Patent No. 6426211
/ GENERAL INFORMATION:
/ APPLICANT: DE BUYL, ERIC
/ LAHAYE, ANDR E
/ LEBOUX, PIERRE
/ AMORY, ANTOINE
/ DETROZ, REN
/ ANDRE, CHRISTOPHE
/ VETTER, ROMAN
/ TITLE OF INVENTION: XYLANASE DERIVED FROM A BACILLUS SPECIES,
/ EXPRESSION VECTORS FOR SUCH XYLANASE AND
/ OTHER PROTEINS, HOST ORGANISMS THEREFOR AND
/ USE THEREOF
/ NUMBER OF SEQUENCES: 35
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: WILLIAN BRINKS HOFER GILSON & LIONE, P.C.
/ STREET: 2000 K St., N.W., Suite 200
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20006
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/073,055
/ FILING DATE: 05-May-1998
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/275,526
/ FILING DATE: 15-JUL-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Gadiano, Wilhem F.
/ REGISTRATION NUMBER: 37,136
/ REFERENCE/DOCKET NUMBER: 4121-49

```

```

/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 429-0625
/ TELEFAX: (202) 293-0625
/ TELEX: 650 383 5605
/ INFORMATION FOR SEQ ID NO: 35:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1022 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: genomic DNA
/ SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-073-055-35

Query Match          99.5%; Score 184; DB 4; Length 1022;
Best Local Similarity 100.0%; Pred. No. 1.2e-51;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATGTAACGCTTGTATCTATTCATTGTTATCAAGGATTTATACAAACAAAGAGA 60
Db 1 TCATGTAACGCTTGTATCTATTCATTGTTATCAAGGATTTATACAAACAAAGAGA 60
QY 61 CATCCATGCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAGGCTGCATGAAGGA 120
Db 61 CATCCATGCGGGTTAAAGCAGTATCGTTCATCTAACAGAGAGGCTGCATGAAGGA 120
QY 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAAACATACGGATGAAAAAGGAGGGAT 180
Db 121 GGTGATGGGTTTTTCATCTTAGGGATGACAGAAACATACGGATGAAAAAGGAGGGAT 180
QY 181 GGAAA 185
Db 181 GGAAA 185

```

```

RESULT 11
US-08-232-463-14/c
/ Sequence 14, Application US/08232463
/ Patent No. 5670367
/ GENERAL INFORMATION:
/ APPLICANT: DORNER, F.
/ APPLICANT: SCHEIFLINGER, F.
/ APPLICANT: FALKNER, F. G.
/ TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
/ NUMBER OF SEQUENCES: 52
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Foley & Lardner
/ STREET: 1800 Diagonal Road, Suite 500
/ CITY: Alexandria
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22313-0299
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/232,463
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/935,313
/ FILING DATE:
/ APPLICATION NUMBER: EP 91 114 300.6
/ FILING DATE: 26-AUG-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: 30472/114 IMMU
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703)836-9300
/ TELEFAX: (703)683-4109

```



TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZ9pt-Fls  
US-08-232-463-14

Query Match 19.4%; Score 35.8; DB 1; Length 7218;  
Best Local Similarity 15.2%; Pred. No. 0.059;  
Matches 25; Conservative 78; Mismatches 61; Indels 0; Gaps 0;

QY 22 ATTTCATTTGTATCAAGAGTTTATACACAAAGAGAGATCCATGCCGGGTTAAAGCA 81  
Db 1480 ATTACCTATCTATGCAAGTAGTTAAAGAGATAGAGAATTTCGTACRRRRRRRRRR 1421

QY 82 GTATCGTTCATCAAGAGAGGCTGCATGAAGAGGAGTGATGGGTTTTCATCTTA 141  
Db 1420 RRR 1361

QY 142 GGGATCAGACAGAAATACCGATGAAAGAGGAGGATGCGAA 185  
Db 1360 RRR 1317

RESULT 12  
US-08-956-171E-38  
Sequence 38, Application US/0895617E  
Patent No. 6593114  
GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
Gil H. Choi  
Patrick S. Dillon  
Craig A. Rosen  
Steven C. Barash  
Michael R. Fannon  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5256  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS, version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,171E  
FILING DATE: 20-Oct-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/009,861  
FILING DATE: January 5, 1996  
APPLICATION NUMBER: 08/781,986  
FILING DATE: January 3, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark J. Hyman  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB248P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (240) 314-1224  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23439 base pairs  
TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 38:  
US-08-956-171E-38

Query Match 17.4%; Score 32.2; DB 4; Length 23439;  
Best Local Similarity 52.2%; Pred. No. 1.5;  
Matches 70; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 44 TATACACAAACAGAGACATCCATCCCGGTTAAAGCAGTATCGTTCATCTAACAGAGA 103  
Db 11481 TATGTATATAACGTGTGATGGACTTAGCGGTTCAATGTTTCAGTCGTTGGTGGCGATAC 11540

QY 104 AGGCTGTCATGAAGGAGGTGATGGGTTTTTCATCTTAGGAGATGACAGAAACAATACGGAT 163  
Db 11541 ACGACCAAAAGAAATTACGAACGTTTGTTCAAATTATGGAAGATTAATGATACCGAT 11600

QY 164 GAAAGAGGAGG 177  
Db 11601 AAAAAACGAGATG 11614

RESULT 13  
US-08-781-986A-38  
Sequence 38, Application US/08781986A  
Patent No. 6737248  
GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5255  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/781,986A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benson, Bob  
REGISTRATION NUMBER: 30,446  
REFERENCE/DOCKET NUMBER: PB248PP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23439 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-781-986A-38

Query Match 17.4%; Score 32.2; DB 4; Length 23439;  
Best Local Similarity 52.2%; Pred. No. 1.5;  
Matches 70; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 44 TATACACAAACAGAGACATCCATCCCGGTTAAAGCAGTATCGTTCATCTAACAGAGA 103  
Db 11481 TATGTATATAACGTGTGATGGACTTAGCGGTTCAATGTTTCAGTCGTTGGTGGCGATAC 11540

QY 104 AGGCTGTCATGAAGGAGGTGATGGGTTTTTCATCTTAGGAGATGACAGAAACAATACCGAT 163

```
Db 11541 ACACCAAGAAATACGAACGTTTCTTCAATTAAGGAGATTAAATGATACGAT 11600
QY 164 GAAAAAAGGAGAGG 177
Db 11601 AAAAAAACGAGATG 11614

RESULT 14
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
```

```
NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664855)..(1664855)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
```

```
Query Match 17.0%; Score 31.4; DB 4; Length 1664976;
Best Local Similarity 59.6%; Pred. No. 17;
Matches 53; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 17 GATCTATTTTCATTTGATCAAGGATTTATACAAACAGAGACATCCATGCCGGGTTA 76
Db 492343 GAATCATGTCGATTGAATCGTAGTTTATCATCATTTAAGCTACTTACAAGCCTTTCTA 492402
QY 77 AGCAGTATCGTTCCATCTAACAGAGAAG 105
Db 492403 ATTCAGATTTTCTTTTCAACTGGAAG 492431
```

```
RESULT 15
US-09-692-570-1
Sequence 1, Application US/09692570
Patent No. 6797466
GENERAL INFORMATION:
APPLICANT: Bult et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
Patent No. 6797466
FILE REFERENCE: PB275C1
CURRENT APPLICATION NUMBER: US/09/692,570
CURRENT FILING DATE: 2003-01-14
PRIORITY APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
PRIOR APPLICATION NUMBER: US 08/916,421
PRIOR FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patentin version 3.1
SEQ ID NO 1
```

```
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc_feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10398)..(10398)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
```

```
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (779676)..(779676)
```

```
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
```

Query Match 17.0%; Score 31.4; DB 4; Length 1664976;  
Best Local Similarity 59.6%; Pred. No. 17;  
Matches 53; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

```
Qy 17 GATCTATTTCATTTGTATCAAGGATTATACACAAAGACATCCATCGCGGTGA 76
Db 492343 GAATCATGTCGATTGATCGTAGGTTTATCATCATTTAAGCTACTTACAGCCTTCTA 492402

Qy 77 AAGCAGTATCGTTCCATCTAACAGAGAAG 105
Db 492403 ATTCAGATTTTCTTTTCTTCAACTGAGAAG 492431
```

Search completed: November 11, 2004, 01:34:27  
Job time : 25.2524 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 21:26:43 ; Search time 118.282 Seconds  
(without alignments)  
8421.106 Million cell updates/sec

Title: US-09-909-207-26

Perfect score: 185

Sequence: 1 TCATGTAACCTGCCTTGATC.....AAAAGAGAGGAGGATGAAA 185

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3611042 seqs, 2692057975 residues

Total number of hits satisfying chosen parameters: 7222084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	184	99.5	185	9	US-09-909-207-26
2	37	20.0	634	13	Sequence 26, Appl
3	37	20.0	634	15	Sequence 209848,
4	35	18.9	624	13	Sequence 209848,
5	35	18.9	624	15	Sequence 182736,
6	34.8	18.8	1500	15	Sequence 182736,
7	33.2	17.9	260803	17	Sequence 46834, A
8	33	17.8	445	16	Sequence 84, Appl
9	32.8	17.7	355	16	Sequence 8116, Ap
10	32.8	17.7	3045	15	Sequence 141151,
11	32.2	17.4	23439	8	Sequence 41036, A
12	32.2	17.4	23439	16	Sequence 38, Appl

c 13	32	17.3	405	9	US-09-960-352-4132	Sequence 4132, Ap
c 14	31.6	17.1	1079	13	US-10-027-632-117073	Sequence 117073,
c 15	31.6	17.1	1079	15	US-10-027-632-117073	Sequence 117073,
c 16	31.6	17.1	684973	9	US-09-263-959-1	Sequence 1, Appl
c 17	31.2	16.9	326	18	US-10-425-115-72247	Sequence 72247, A
c 18	31.2	16.9	497	16	US-10-424-599-115848	Sequence 115848,
c 19	31	16.8	390	9	US-09-736-457-1223	Sequence 1223, Ap
c 20	31	16.8	390	9	US-09-902-941-1223	Sequence 1223, Ap
c 21	31	16.8	390	9	US-09-849-626-1223	Sequence 1223, Ap
c 22	31	16.8	390	14	US-10-017-754-1223	Sequence 1223, Ap
c 23	31	16.8	390	15	US-10-113-872-1223	Sequence 1223, Ap
c 24	31	16.8	390	15	US-10-283-017-1223	Sequence 1223, Ap
c 25	31	16.8	32183	9	US-09-764-869-1494	Sequence 1494, Ap
c 26	31	16.8	32183	14	US-10-091-504-1494	Sequence 1494, Ap
c 27	31	16.8	32183	16	US-10-227-577-1494	Sequence 1494, Ap
c 28	31	16.8	301692	16	US-10-428-487-195	Sequence 195, Appl
c 29	31	16.8	310268	17	US-10-367-094-195	Sequence 128655,
c 30	30.8	16.6	804	13	US-10-027-632-128655	Sequence 128655,
c 31	30.8	16.6	804	15	US-10-027-632-128655	Sequence 3452, Ap
c 32	30.8	16.6	1132	9	US-09-938-842A-3452	Sequence 3452, Ap
c 33	30.8	16.6	1132	11	US-09-938-842A-3452	Sequence 496, App
c 34	30.6	16.5	58723	13	US-10-087-192-496	Sequence 16465, A
c 35	30.2	16.3	930	16	US-10-282-122A-16465	Sequence 282580,
c 36	30	16.2	561	13	US-10-027-632-282580	Sequence 282580,
c 37	30	16.2	561	15	US-10-027-632-282580	Sequence 264187,
c 38	30	16.2	831	16	US-10-305-720-18	Sequence 264187,
c 39	29.8	16.1	646	13	US-10-027-632-264187	Sequence 53, Appl
c 40	29.8	16.1	646	15	US-10-027-632-264187	Sequence 6389, Ap
c 41	29.8	16.1	178896	17	US-10-450-826-53	Sequence 478, App
c 42	29.6	16.0	422	18	US-10-674-124A-6389	Sequence 948, App
c 43	29.6	16.0	610	14	US-10-116-802-478	Sequence 49, Appl
c 44	29.6	16.0	977	10	US-09-764-872-948	
c 45	29.6	16.0	1143	9	US-09-738-626-49	

## ALIGNMENTS

### RESULT 1

US-09-909-207-26

; Sequence 26, Application US/09909207

; Patent No. US20020115181A1

; GENERAL INFORMATION:

; APPLICANT: ANDREE LAHAYE

; ERIC DE BUYL

; PIERRE LEDOUX

; RENE DETROZ

; TITLE OF INVENTION: Xylanase, microorganisms produced it,  
DNA molecule, processes for preparation of this xylanase  
and uses thereof

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSER: WILLIAN BRINKS HOFER GILSON & LIONE

; STREET: 2000 K St., N.W., Suite 200

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20006

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA: US/09/909,207

; APPLICATION NUMBER: 19-Jul-2001

; FILING DATE: 19-Jul-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/470,953

; FILING DATE: 06-JUNE-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Wilhem F. Gadiano, Esq.

; REGISTRATION NUMBER: 37,136

```
; REFERENCE/DOCKET NUMBER: 4121-40
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-429-0625
; TELEFAX: (202) 293-1850
; TELEX: 650 383-5605
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 185 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-909-207-26
```

```
Query Match          99.5%; Score 184; DB 9; Length 185;
Best Local Similarity 100.0%; Pred. No. 6.4e-49;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATCTAACTCGCCTTGATCTATTTTCATTCATCAAGGATTTATACAAACAAGAGA 60
DB 1 TCATCTAACTCGCCTTGATCTATTTTCATTCATCAAGGATTTATACAAACAAGAGA 60

QY 61 CATCATCGCGGTTAAAGCAGTATCGTTCCATCTAACAGAGAAGGCTGCATGAAAGGA 120
DB 61 CATCATCGCGGTTAAAGCAGTATCGTTCCATCTAACAGAGAAGGCTGCATGAAAGGA 120

QY 121 GGTGATGGTTTTTCATCTTAGGATGACAGAACTACCGATGAAAAAGAGAGGAT 180
DB 121 GGTGATGGTTTTTCATCTTAGGATGACAGAACTACCGATGAAAAAGAGAGGAT 180

QY 181 CGAAA 185
DB 181 GGA AAA 185
```

```
RESULT 2
US-10-027-632-209848/c
; Sequence 209848, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 209848
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-209848
```

```
Query Match          20.0%; Score 37; DB 15; Length 634;
Best Local Similarity 52.7%; Pred. No. 0.18;
Matches 79; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 36 AAAGGATTTATACAAACAAGAGACATCCATGCGGGTTAAAGCAGTATCGTTCCATCT 95
DB 250 AAAGGATCAACAATAAACAGATGACCTACAGACAGGTGATATATTAGCAATCAACTT 191

QY 96 AACAGAGAAGGCTGCATGAAAGGAGGTGATGGGTTTTTCATCTTAGGGATGACAGACA 155
DB 190 ATCTAACBAGGAATTAATACTGGAATATATAAGTAATTCAAACATCTCAATAAAAAATA 131

QY 156 ATACGGATGAAAAAGGAGGATGGA 185
DB 130 ATTTGATTAATAATGCGTAAAGAATTGAA 101

RESULT 4
US-10-027-632-182736/c
; Sequence 182736, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
```

```
Query Match          20.0%; Score 37; DB 13; Length 634;
Best Local Similarity 52.7%; Pred. No. 0.18;
Matches 79; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
```

```

; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 182736
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-182736

```

```

Query Match      18.9%; Score 35; DB 13; Length 624;
Best Local Similarity 51.3%; Pred. No. 0.79;
Matches 80; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 18 ATCTATTTTCATTTGTTATCAAGAGTTTATACAAACAGAGACATCCATGCCGGTTAA 77
DB 503 ATCAATATGTTTGAATGAGAAAGAGCTACCAATTAACCAAGAGACAAATAGACAGGAAG 444
QY 78 AGCAGTATCGTTCATCTTAACAGAGAGGNCCTGATGAAGAGGAGTGGTGGTTTTCAT 137
DB 443 AGCAGTATCATGAGACTGATGGAAGAAAGCTTAAGGAAAGACTGATGAATAGTATTGAA 384
QY 138 CTTAGGATGACAGAAACATACGGATGAAAAAGGA 173
DB 383 GTTAGAAATTTACAAAAATTAACAAATTAAGTGA 348

```

```

RESULT 5
US-10-027-632-182736/c
; Sequence 182736, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 182736
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-182736

```

```

Query Match      18.9%; Score 35; DB 15; Length 624;
Best Local Similarity 51.3%; Pred. No. 0.79;
Matches 80; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 18 ATCTATTTTCATTTGTTATCAAGAGTTTATACAAACAGAGACATCCATGCCGGTTAA 77
DB 503 ATCAATATGTTTGAATGAGAAAGAGCTACCAATTAACCAAGAGACAAATAGACAGGAAG 444
QY 78 AGCAGTATCGTTCATCTTAACAGAGAGGNCCTGATGAAGAGGAGTGGTGGTTTTCAT 137
DB 443 AGCAGTATCATGAGACTGATGGAAGAAAGCTTAAGGAAAGACTGATGAATAGTATTGAA 384
QY 138 CTTAGGATGACAGAAACATACGGATGAAAAAGGA 173
DB 383 GTTAGAAATTTACAAAAATTAACAAATTAAGTGA 348

```

```

RESULT 6
US-10-369-493-46834
; Sequence 46834, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 46834
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-10-369-493-46834

```

```

Query Match      18.8%; Score 34.8; DB 15; Length 1500;
Best Local Similarity 65.4%; Pred. No. 1.3;
Matches 51; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 20 CTATTTTCATTTGTTATCAAGAGTTTATACAAACAGAGACATCCATGCCGGTTAAAG 79
DB 1423 CTATTTGATTTGTTATGAAGAACATTTACACTCACAAAGGATCTCAATACAGCTTTGAAG 1482
QY 80 CAGTATCGTTCATCTAA 97
DB 1483 AGCTTCGAAAAAACTAA 1500

```

```

RESULT 7
US-10-388-838-84/c
; Sequence 84, Application US/10388838
; Publication No. US20040180344A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001600
; CURRENT APPLICATION NUMBER: US/10/388,838
; CURRENT FILING DATE: 2003-03-14
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 260803
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

```

```

; LOCATION: (1)...(260803)
; OTHER INFORMATION: n = A,T,C or G
US-10-388-838-84

Query Match
Best Local Similarity 17.9%; Score 33.2; DB 17; Length 260803;
Matches 89; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 2 CATGTAACTCGCTTGATCTATTTCATTTGTATCAAGGATTATACACAACAGAGAC 61
Db 183802 CATGTAAATGTTAGAGTCTGTAAGTTTGTAGCAGTTTGTATGCGCAGCACAGAAAAC 183743

QY 62 ATCCATCGCGGTTAAAGCAGTATCGTCCATCTAACAGAGAGNCTGCATGAAGGAG 121
Db 183742 TACCATAATTTGGTACAGATTTACCATCTCAACTAAGAAATCAGAAAAATACATTAA 183683

QY 122 GTGATGGGTTTTTCATCTTTAGGATGACAGAAACAATACGGATGAAAAAGGAGGATG 181
Db 183682 ATTATGGCTTTCTGATCTTTGGACAACAGGACGACAGGAGAGCAATCTAAGAGAGAGT 183623

QY 182 GAA 184
Db 183622 GAA 183620

RESULT 8
US-10-424-599-8116/c
; Sequence 8116, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 8116
; LENGTH: 445
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_107339C.1
US-10-424-599-8116

Query Match
Best Local Similarity 17.8%; Score 33; DB 16; Length 445;
Matches 78; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 20 CTATTTTCATTTGTATCAAGGATTATACACAACAGAGACATCCATGCCGGGTTAAAG 79
Db 245 CAATTTTCATATAAAGAATCATTTATCTAAATACGATTTTGTCTTTTATTTGCTTAAAA 186

QY 80 CAGTATCGTTTCATCTAACAGAGAGGCTGCATGAAGAGGATGATGGGTTTTTCATCT 139
Db 185 CATTTTCTCTTTTGAACAATCTGAATAGTATGTTAAATATGATTTTAGTTTTTTTCT 126

QY 140 TAGGATGACAGAACAAATACGGATGAAAAAGGA 173
Db 125 TCTTTTCTTAATGAATATATATTTAAATATGA 92

RESULT 9
US-10-424-599-141151/c
; Sequence 141151, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua

```

```

; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 141151
; LENGTH: 355
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_98471C.1
US-10-424-599-141151

Query Match
Best Local Similarity 17.7%; Score 32.8; DB 16; Length 355;
Matches 55; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 91 CATCTAACAGAGAGGCTGCATGAAGAGGATGATGGTTTTCATCTTAGGGATGACA 150
Db 329 CATTTCCAAAAGGATGAACAAATGGAAGGAGATGATGATATTGGTTTTCAGCTGTACT 270

QY 151 GAACAATACGGATGAAAAAAGGAGGCGGATGGA 183
Db 269 GAACAATACAGAGAGGCGGGAAGATGGA 237

RESULT 10
US-10-369-493-41036
; Sequence 41036, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 41036
; LENGTH: 3045
; TYPE: DNA
; ORGANISM: Bacillus halodurans
US-10-369-493-41036

Query Match
Best Local Similarity 17.7%; Score 32.8; DB 15; Length 3045;
Matches 76; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 27 ATTGTGATCAAGAGATTATACACAACAGAGACATCCATGCCGGTTAAAGCATC 86
Db 1496 ATATGGAAGTCGATGTATGCAAAACCAAAAGACATTTAGGACTATTTHACGACGATC 1555

QY 87 GTTCCATCTAACAGAGAGGCTGCATGAAGAGGATGATGGGTTTTTTCATCTTAGGGAT 146
Db 1556 CGAAAAAGCCGTACATTAGCTGTGAATATATGATCGATGGGCAACTCTTTTAGTGGGA 1615

QY 147 GACAGAACATACGGATGAAAAAGGAGA 175
Db 1616 TGCAACAAGTATCTGAGTTGGAACAGAAA 1644

RESULT 11
US-08-781-986A-38
; Sequence 38, Application US/08781986A
; Publication No. US20030054436A1

```



GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5255  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/781,986A  
FILING DATE: 27-Dec-2002  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/956,171  
FILING DATE: October 20, 1997  
APPLICATION NUMBER: 60/009,861  
FILING DATE: January 5, 1996  
APPLICATION NUMBER: 08/781,986  
FILING DATE: January 3, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark J. Hyman  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB248PID1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (240) 314-1224  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23439 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-781-986A-38

Query Match 17.4%; Score 32.2; DB 8; Length 23439;  
Best Local Similarity 52.2%; Pred. No. 27;  
Matches 70; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 44 TATACACAAAGAGACATCCATCGCGGTTAAAGCAGTATCGTTCCATCTAACAGAGA 103  
DB 11481 TATGTATAAACGCTGTGATGGACTTAGCGGTTTCATCTAGTCTGGTGGCGATAAC 11540

QY 104 AGCNCCTGCATGAAAGAGGTGATGGGTTTTTTCATCTTAGGGATGACAGAACATACGGAT 163  
DB 11541 ACGACCAGAAAGAAATACGAAACGTTTGTTCAAATTATGGAAGATTAAATGATAACGAT 11600

QY 164 GAAAAAGGAGG 177  
DB 11601 AAAAAACGAGATG 11614

RESULT 12  
US-10-329-624-38  
Publication 38, Application US/10329624  
Publication No. US20040043037A1  
GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
Gil H. Choi  
Patrick S. Dillon  
Craig A. Rosen  
Steven C. Barash  
Michael R. Fannon  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5256  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland

COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/329,624  
FILING DATE: 27-Dec-2002  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/956,171  
FILING DATE: October 20, 1997  
APPLICATION NUMBER: 60/009,861  
FILING DATE: January 5, 1996  
APPLICATION NUMBER: 08/781,986  
FILING DATE: January 3, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark J. Hyman  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB248PID1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (240) 314-1224  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23439 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-10-329-624-38

Query Match 17.4%; Score 32.2; DB 16; Length 23439;  
Best Local Similarity 52.2%; Pred. No. 27;  
Matches 70; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 44 TATACACAAAGAGACATCCATCGCGGTTAAAGCAGTATCGTTCCATCTAACAGAGA 103  
DB 11481 TATGTATAAACGCTGTGATGGACTTAGCGGTTTCATCTAGTCTGGTGGCGATAAC 11540

QY 104 AGCNCCTGCATGAAAGAGGTGATGGGTTTTTTCATCTTAGGGATGACAGAACATACGGAT 163  
DB 11541 ACGACCAGAAAGAAATACGAAACGTTTGTTCAAATTATGGAAGATTAAATGATAACGAT 11600

QY 164 GAAAAAGGAGG 177  
DB 11601 AAAAAACGAGATG 11614

RESULT 13  
US-09-960-352-4132/c  
Sequence 4132, Application US/09960352  
Patent No. US20020137139A1  
GENERAL INFORMATION:  
APPLICANT: Warren, Wesley C.  
APPLICANT: Tao, Nengbing  
APPLICANT: Byatt, John C.  
APPLICANT: Mathialagan, Nagappan  
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION  
FILE REFERENCE: 16511.006/37-21 (10298)C  
CURRENT APPLICATION NUMBER: US/09/960,352  
CURRENT FILING DATE: 2001-09-24  
NUMBER OF SEQ ID NOS: 15112  
SEQ ID NO 4132  
LENGTH: 405  
TYPE: DNA  
ORGANISM: Bos taurus  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (372)  
OTHER INFORMATION: unsure at all n locations

